Pattern Recognition Employing Spatially Variant Unconstrained Correlation Filters

By

Akber Abid Gardezi

SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY AT THE UNIVERSITY OF SUSSEX

School of Engineering and Informatics

University of Sussex

Brighton , UK

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature

Akber Abid Gardezi

Dated: 12th September 2013

Acknowledgements

Firstly before anything I would like to thank my parents for always being there for me and supporting me in every way humanly possible. I am deeply thankful to my sisters for their love and support all throughout my life. I am especially grateful to Amna my fiancé for bearing up with me all these years. Also to all of my extended family members who have always prayed for my success during my stay at University of Sussex.

I am deeply grateful to my supervisors Dr. Rupert Young , Prof. Chris Chatwin and Dr. Phil Birch for their kind support throughout my degree. Going well beyond the call of duty to assist me in every way they could and to facilitate me in times of hardships. No matter how much I thank them wouldn't be enough.

Secondly I would like to thank my dear friend and colleague Dr. Bhargav Mitra for his kind help and support academically and emotionally and being a beacon of motivation for me to push forward when times were bleak. Also my dear friend Ahmed Al-Kandri truly an amazing man and a great source of motivation for me.

Also I would like to thank Bushra Hassan for being a kind friend and kindred and loving soul and a source of great support during the critical time of my degree. I would like to thank my dear friends Tabassam Qureshi and Rana Bilal for always being there whenever I need them and always ready to help with no second thoughts without them this journey wouldn't have been possible. Also a great thank you to Abu Ali for being a great friend and a younger brother I never had.

Lastly but not the least I would like to thank Waheed Bhai and Auntie Nergis with whom I have lived in Brighton for five wonderful years, thank you for being my family here in the UK and my depth of gratitude towards you have no bounds.

Lastly a big thanks to all of my research colleagues at Sussex University for always being there and helpful.

Abstract

A spatial domain Optimal Trade-off Maximum Average Correlation Height (SPOT-MACH) filter is proposed in this thesis. The proposed technique uses a pre-defined fixed size kernel rather than using estimation techniques. The spatial domain implementation of OT-MACH offers the advantage that it does not have shift invariance imposed on it as the kernel can be modified depending upon its position within the input image. This allows normalization of the kernel and allows inclusion of a space domain non-linearity to improve performance.

The proposed SPOT-MACH filter can be used to maximize the height of the correlation peak in the presence of distortions of the training object and provide resistance to background clutter. One of the major characteristics of the SPOT-MACH filter is that it can be tuned to maximize the height and sharpness of the correlation peak by using trade-offs between distortion tolerance, peak sharpness and the ability to suppress clutter noise.

A number of non-parametric local regression techniques offer a simplified approach to pattern recognition problems which employ linear filtering using low pass filters designed using moving window local approximations. In most of these cases the algorithms search for a region of interest near the point of estimation for various prevailing conditions which fit the required criteria. These estimates are calculated for a defined window size which is determined as being the largest area within which the estimators do not widely vary from the criteria. The only drawback in this approach is that the window size is directly proportional to the required computational resources and would adversely affect the performance of the system if the moving window size is not proportionate to the resources.

The proposed filter employs an optimization technique using low-pass filtering to highlight the potential region of interests in the image and then restricts the movement of the kernel to these regions to allow target identification and to use less computational resources. Also another optimization technique is also proposed which is based on an entropy filter which measures the degree of randomness between two changing scenes and would return the area where change has occurred i.e. the target object might be present. This approach gives a more accurate region of interest than the low-pass filtering approach.

Apart from the software based optimization approaches two hardware based enhancement techniques have also been proposed in this thesis. One of the approaches employs Field

Programmable Gate Array (FPGA) to perform correlation process employing the inbuilt multipliers and look up tables and the other one uses Graphical Processing Unit (GPU) to do parallel processing of the input scene.

Also in this thesis a detailed analysis of SPOT-MACH has been carried out by comparing with popular feature based techniques like Scale Invariant Feature Transform (SIFT) and a comparison matrix has been created.

The proposed filter uses a two-staged approach using speed optimizations and then detection of targets from input scenes. Both visible and Forward Looking Infrared (FLIR) imagery data sets have been used to test the performance of filter.

Contents

List of Figures

List of Tables

List of Publications

The publications resulting from this thesis are presented:

1. Birch, P., Gardezi, A., Mitra, B., Young, R., Chatwin, C., "An optical space domain volume holographic correlator" in *Optical Pattern Recognition XX, Orlando ,USA,2009*

2. Gardezi, A., Birch, P., Kypraios, I., Young, R., Chatwin, C., "Implementation of the Maximum Average Correlation Height (MACH) filter in the spatial domain for object recognition from clutter backgrounds" in *Automatic Target Recognition XX; Acquisition, Tracking, Pointing, and Laser Systems Technologies XXIV; and Optical Pattern Recognition XXI , Orlando ,USA,2010*

3. Birch, P. M., Gardezi, A., Young, R., Chatwin, C., "Volume holographic MACH correlator" in *Automatic Target Recognition XX; Acquisition, Tracking, Pointing, and Laser Systems Technologies XXIV; and Optical Pattern Recognition XXI, Orlando ,USA,2010*

4. Mitra, B. K., Hassan, W., Birch, P., Gardezi, A., Young, R., Chatwin, C., "A two-stage approach to detect abandoned baggage in public places" in *Visual Information Processing XIX , Orlando ,USA,2010*

5. Gardezi, A., Alkandri, A., Birch, P., Young, R., Chatwin, C., "A space variant maximum average correlation height (MACH) filter for object recognition in real time thermal images for security applications" in *Optics and Photonics for Counterterrorism and Crime Fighting VI and Optical Materials in Defence Systems Technology VII , Toulouse, France , 2011*

6. Gardezi, A.; Alkandri, A.; Birch, P.; Qureshi, T.; Young, R.; Chatwin, C.; , "An implementation and performance evaluation of a space variant OT-MACH filter for a security detection application using FLIR sensor," *Applied Imagery Pattern Recognition Workshop (AIPR), 2010 IEEE 39th* , vol., no., pp.1-7, 13-15 ,Washington D.C. USA Oct. 2010

7. Gardezi, A., Al-Kandri, A., Birch, P., Young, R., Chatwin, C., "Enhancement of the speed of space-variant correlation filter implementations by using low-pass pre-filtering for kernel placement and applications to real-time security monitoring" in *Optical Pattern Recognition XXII, Prague , CZECH Republic , 2011*

8. Alkandri, A., Gardezi, A., Birch, P., Young, R., Chatwin, C., "Parameter optimization of the optimal trade-off maximum average correlation height filter for FLIR imaging in high clutter environments" in *Optical Pattern Recognition XXII ,Prague , CZECH Republic , 2011*

9. Akber Gardezi ; Nagachetan Bangalore ; Ahmed Al-Kandri ; Philip Birch ; Rupert Young, et al. "Application of speed-enhanced spatial domain correlation filters for real-time security monitoring", Proc. SPIE 8189, Optics and Photonics for Counterterrorism and Crime Fighting VII; Optical Materials in Defence Systems Technology VIII; and Quantum-Physics-based Information Security, Baltimore , U.S.A , 2012

10. Alkandri, A., Gardezi, A., Bangalore, N., Birch, P., Young, R., Chatwin, C., "Automatic parameter adjustment of difference of Gaussian (DoG) filter to improve OT-MACH filter performance for target recognition applications" in Electro-Optical and Infrared Systems: Technology and Applications VIII, Baltimore , U.S.A , 2012

List of Symbols and Acronyms

ς	Coefficient of Correlation
$\Im'$	Two dimensional Fourier Transform
$\lvert R(o,v)\rvert$	Amplitude of Fourier Transform
$\Phi(o,v)$	Phase of Fourier Transform
σ	Standard deviation of the correlation plane
τ	Conjugate Transpose
u_{opt}	Optimal Shape Vector
∇_u	Gradient with respect to zero
S_x	Similarity Matrix
α	OT-MACH Filter Parameter
β	OT-MACH Filter Parameter
γ	OT-MACH Filter parameter
D_x	Diagonal Power Spectrum density
u_{ot}	Optimal Shape Vector for EMACH
λ	Eigen Value

ECP	Equal Correlation Peak
IPR	Intelligent Pattern Recognition
MVSDF	Minimum Variance Synthetic Discriminant Function
MACE	Minimum Average Correlation Energy
MACH	Maximum Average Correlation Height
ASM	Average Similarity Measure
ACH	Average Correlation Height
MSE	Mean Squared Error
OT-MACH	Optimal Trade-off Maximum Average Correlation Height
MASM	Modified Average Similarity Measure
SPOT-MACH	Spatial Domain OT-MACH
ODOG	Oriented Difference-o- Gaussian
SCORR	Spatial Correlator
SIFT	Scale Invariant Feature Transform
ONV	Output Noise Variance
EMACH	Extended MACH
AICH	All image correlation height
COPI	Correlation Output peak intensity
PCE	Peak to Correlation Energy
PSR	Peak to Side-lobe Ratio
SDF	Synthetic Discriminant Function
FLIR	Forward Looking Infra Red
FFT	Fast Fourier Transform
LUT	Look up Table
ASMBL	Advanced Silicon Modular Block
CPU	Central Processing Unit
RAM	Random Access Memory
IP	Intellectual Property

Chapter 1

Introduction

1.1 Overview

Correlation filters are an integral component of many intelligent image pattern recognition (IPR) systems capable of pre-processing data and defining decision making criteria. The field of pattern recognition is based on the principle of detection and identification of desired targets from an unknown input scene where the target may or may not be present. In the presence of the desired target the main objective is to determine its spatial location within the input scene. IPR is a rapidly growing area of research attracting cross disciplinary attention and involvement to develop new technologies.

Some of the major applications of IPR systems include computer vision, facial recognition, bio-metric identification, character recognition, medical diagnosis, video analytics and surveillance etc.

An IPR pattern system operates on the basic principle that the input is an image and the output is a decision signal based on the characteristic features of the input image. However, the problem in this case is that the numbers of available features is usually fewer than those required to describe the actual target object, thus degrading or even preventing its potentially accurate identification [1] , [2].

The characteristics of a general pattern recognition system are different when compared to a statistical hypothesis test where the sensed data is used to decide whether or not to reject a null hypothesis in favour of an alternative hypothesis based on a critical value [3], [4].

In order to better understand the concept of IPR it may be compared with associative memory where the input is considered as a pattern and another pattern is emitted at the

output which is a representative of a general group. The pattern emitted at the output results in reducing the information to some extent but is still not comparable to the reduction which occurs in an IPR system [5].

Historically there are two major approaches used in pattern recognition systems which are known as Statistical and Syntactic pattern recognition. In pattern recognition there is generally an underlying and quantifiable statistical basis or a structure of a pattern which provides the fundamental information about that pattern. Sometimes the pattern can be described by a combination of both of these approaches [6] , [7].

A system based on statistical pattern recognition approach, assumes a statistical basis for any classification algorithm. A set of characteristic measurements and features is extracted from the input data and is used to assign each feature vector to one of the classes. Features are assumed to be generated by a state of nature and therefore the underlying model is a particular probability density function [8] , [9].

In syntactic pattern recognition the absence or presence of a set of features is not considered but the interrelationship or interconnections of features provides important structural information which facilitates classification. However, in this case, it is necessary to quantify and extract structural information in order to assess similarity of patterns [8] , [9].

From this it can be deduced that the integral criterion in an IPR system is decision making which is achieved by an effective information reduction process where it is not possible to reconstruct the input pattern but, nevertheless, to give an accurate decision. This effective information reduction process can be classified as filtering. So at the heart of every IPR system is a filter which does the filtering, i.e. the information reduction, and after that, the correlation characteristics are used to make the required decision. These two steps jointly can be defined as a ‘Correlation Filter ‘in the simplest manner. There are a variety of filters that can be used in both frequency and spatial domains depending upon the scenario; these are discussed in more detail in the forthcoming chapters.

In order to give a better understanding of the addressed research area for this thesis, that is the use of space variant approaches for an effective pattern recognition system which is invariant to background clutter and illumination, the thesis is introduced by answering

the most basic question: ‘What is Correlation?’ and then progresses towards answering the question ‘Why Space Variant Approaches?’.

1.2 What is Correlation?

Correlation can be defined as the process which gives the degree of similarity between two vectors. The correlation approach can not only be expressed in vector form but is also directly applicable to two dimensional images as well. In this case, process of correlation simply involves the moving the center of a filter mask from point to point in an image. At each point the correlation output is the sum of products of the filter coefficients and the corresponding neighborhood pixels in the area covered by the filter mask [10] , [11].

The discrimination criteria for the correlation process between two vectors can be more clearly seen by reference to Figure 1.1.

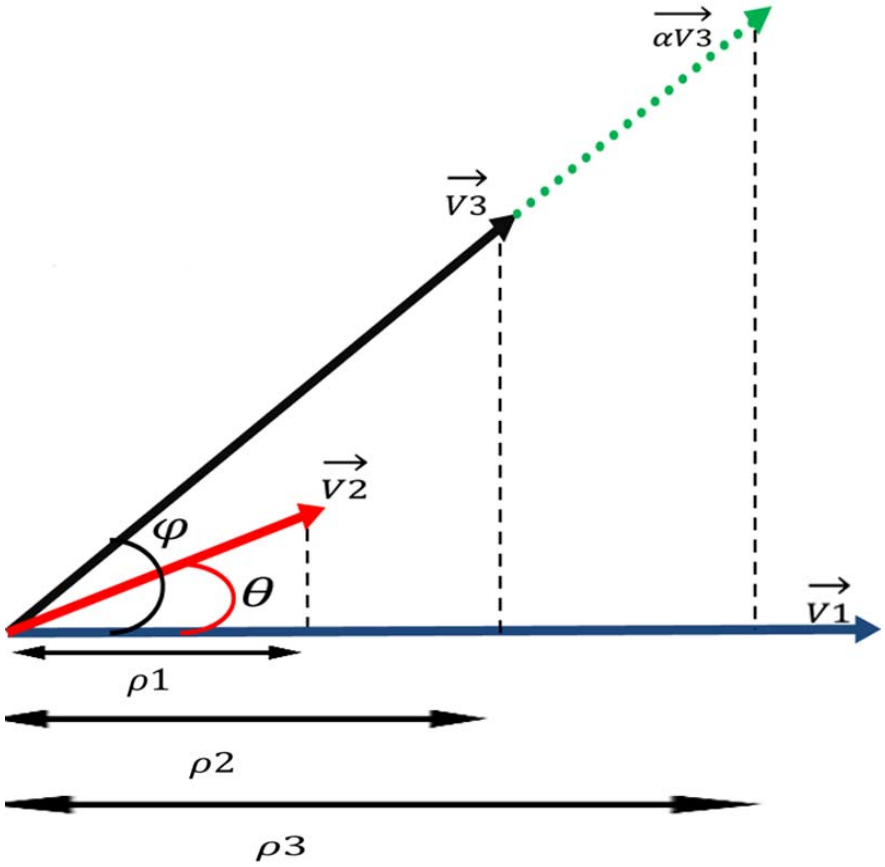


Figure 1.1 Correlation between two vectors $\overrightarrow{V1}$ and $\overrightarrow{V2}$

Initially researchers considered that correlation between two vectors could be measured by taking the projection of one vector on to another vector. However, this may give rise to erroneous results as illustrated in Figure 1.1 [12].

In the example we are trying to measure the correlation between two sets of vectors given as: $\overrightarrow{V1}$, $\overrightarrow{V2}$ and $\overrightarrow{V3}$ and $\overrightarrow{\alpha V3}$. It is apparent from the figure that the vector $\overrightarrow{V3}$ is more correlated to the vector $\overrightarrow{V1}$ as compared to the correlation between $\overrightarrow{V2}$ and $\overrightarrow{V1}$.

However, relative to the magnitude of $\overrightarrow{V1}$, the projection magnitude of $\overrightarrow{V3}$ is larger than the projection amplitude of $\overrightarrow{V2}$. It can also be seen that if vector $\overrightarrow{V3}$ is multiplied by a scalar constant α then the magnitude of the projection of the modified vector on $\overrightarrow{V1}$ increases as well.

There has always been some debate over the issue of what distinguishes an accurate correlation against an erroneous one. Initially researchers stated that if a vector is correlated with another vector and a perpendicular is drawn on the input vector then the greater the distance from the origin, the better is the correlation [12] , [13].

But this hypothesis fails if Figure 1.1 is considered as clearly in this case $\overrightarrow{V3}$ would be the most correlated vector but from Figure 1.1 it is not in line with the input vector $\overrightarrow{V1}$. So when these distance classifier criteria failed researchers developed with a new approach for correlation which involved computing the value of angle θ. This is illustrated by equations 1.1 to 1.3.

$$\overrightarrow{V(i)} \bullet \overrightarrow{V(j)} = \left|\overrightarrow{V(i)}\right| \times \left|\overrightarrow{V(j)}\right| \times \cos\ \theta \tag{1.1}$$

$$\cos\ \theta = \frac{\overrightarrow{V(i)} \bullet \overrightarrow{V(j)}}{\left|\overrightarrow{V(i)}\right| \times \left|\overrightarrow{V(j)}\right|} \tag{1.2}$$

and θ is given by

$$or,\ \theta = \cos^{-1}\left(\frac{\overrightarrow{V(i)} \bullet \overrightarrow{V(j)}}{\left|\overrightarrow{V(i)}\right| \times \left|\overrightarrow{V(j)}\right|}\right) \tag{1.3}$$

By applying a basic trigonometric transformation equation 1.2 is deduced and the discrimination capabilities of the correlator are now dependent upon the value of $\cos\ \theta$. Now if the output value is near to 1 it shows a better correlation since $\overrightarrow{V2}$ is the closest match despite being small in size compared to the other vectors. If the phase of the input vector is aligned with the target vector with its maximum set to 1 then at $\cos\ \theta$=0 where it is equal to 1 there will be maximum correlation and minimum correlation will be $\cos\ \theta$=90. So in this case $\overrightarrow{V2}$ will give the highest correlation peak and a better discrimination criterion which is more robust and accurate [12] , [14].

The correlation approach presented above is in the vector form but for pattern recognition problems a representation to work directly with a two dimensional image is required. In this case a correlation function as shown by equation 1.4 has been created for an image $w(x, y)$ of size $J \times K$ with an image $f(x, y)$ of size $M \times N$ where $J \le M$ and $K \le N$ [10] , [11].

$$c(x,y) = \sum_{s=0}^{J}\sum_{t=0}^{K} f(x,y).w(x+s,y+t) \tag{1.4}$$

For $x = 0,1,2,.....,M-1$ and $y = 0,1,2,.....,N-1$ the summation is computed over the region where w and f overlap [15].

This can be further illustrated by Figure 1.2 which shows the correlation process for a two dimensional image.

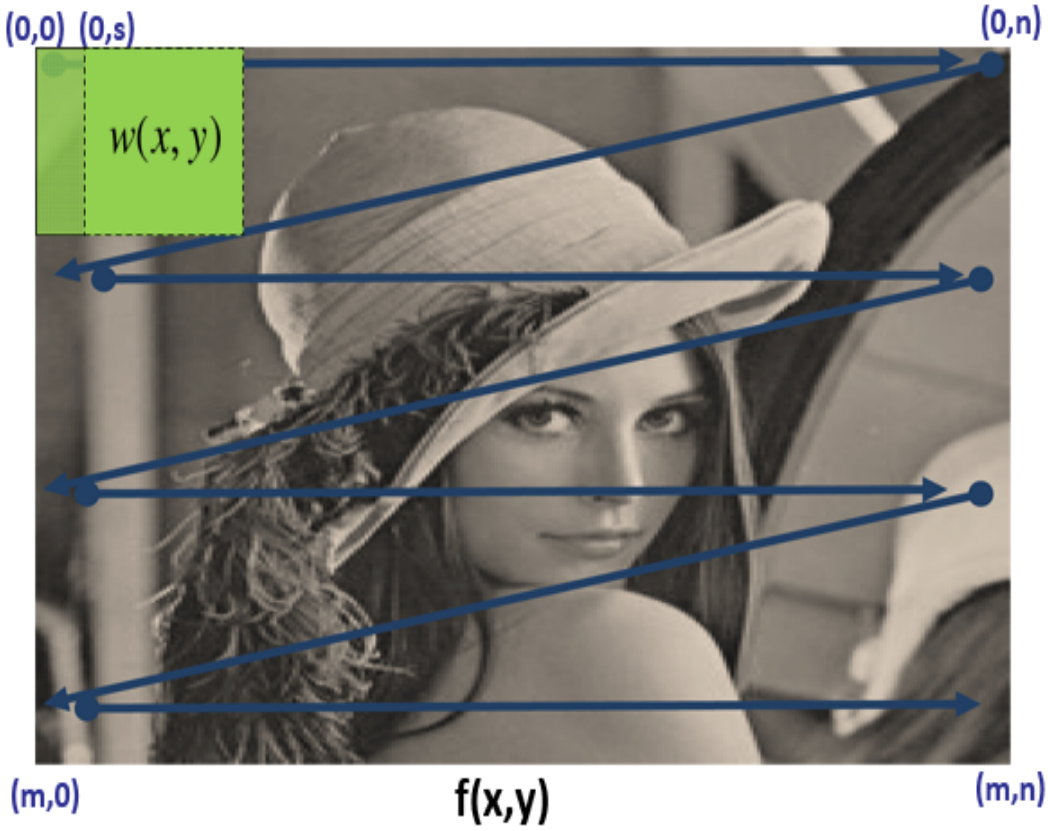


Figure 1.2 Correlation of a two dimensional image [11]

In Figure 1.2 it has been assumed that the origin of f is in the top left corner and the origin of w is at its center. When the coordinates (x, y) are varied w moves around the image producing the function $c(x, y)$. The maximum value of the function c indicates the location where w matches with f. The correlation accuracy is reduced near the edges of the image f. This is due to the correlation function given in equation 1.4 having the drawback of being sensitive to the changes in the amplitude of image f and w as more energy would be locked at the edges hence more chances of false detection at the edges [5], [10] and [16].

This can be overcome by using a correlation coefficient when correlating two images. The correlation coefficient is given in equation 1.5 [10], [11].

$$\varsigma(x,y)=\frac{\sum_s\sum_t\left[f(s,t)-\overline{f}(s,t)\right]\left[w(x+s,y+t)-\overline{w}\right]}{\left\{\sum_s\sum_t\left[f(s,t)-\overline{f}(s,t)\right]^2\sum_s\sum_t\left[w(x+s,y+t)-\overline{w}\right]^2\right\}^{\frac{1}{2}}} \tag{1.5}$$

where $x=0,1,2,.....,M-1$ and $y=0,1,2,.....,N-1$.

Also $\overline{w}$ is the average value of the pixels in w which is evaluated at the beginning with $\overline{f}$ being the average value of f in the region which contains the moving window. The summations are computed in the regions which are overlapping and have common coordinates for f and w. The coefficient of correlation $\varsigma(x,y)$ is scaled between the range of -1 to 1 which is independent to the scale changes in the amplitude of f and w [1], [2] and [11].

Using equation 1.4 gives a normalized approach for correlation which enables pattern matching even in the areas of non-uniform brightness.

1.3 Problem Definition

The main advantage of the correlation function is that it can be normalized for amplitude changes using the correlation coefficient in the spatial domain. But the application of normalisation for changes in size and rotation is quite a computationally intensive task. Due to the computationally intensive nature of spatial correlation we need to normalize for size in the local window which involves spatial scaling and this turns out to be quite a computationally intensive task. In the case the object is oriented at an unknown angle then $w(x,y)$ must be rotated so that it aligns with the degree of rotation of $f(x,y)$ which again is computationally intensive process. Due to the

impracticalities involved in this process, correlation is seldom used in scenarios where unconstrained rotation is present in the target scene [11] , [17].

Due to the computationally intensive nature of correlation, over the years most of the pattern recognition applications have employed alternative ways of computing correlation by the means of using the Fast Fourier Transform (FFT) in the frequency domain [5]. In the case where f and w are the same size, the frequency domain approaches have proved to be more efficient than the direct implementation in the spatial domain. In addition methods have been developed over the last twenty years to train the filter to be invariant to distortions of the target object and most of these methods are implemented in the frequency domain. It was deduced by Gonzalez and Woods that when the number of non-zero terms in w are less than 132, i.e. a local window of size '13 x 13' or less is used then using the spatial domain implementation is more efficient than the FFT approach [10]. This however puts a restraint on the size of local window that can be used efficiently but provides an indication of when FFT approaches should be considered. However, most importantly, the correlation coefficient method cannot be implemented in the frequency domain. Thus the frequency domain approach only implements as un-normalized correlation calculation but with the trade-off of increased speed [11], [18] and [19].

The distortion invariant filters as implemented in the frequency domain are hence un-normalized in nature but allow increased computational speed. But with the rapid advances in the digital field faster and faster computational resources are coming on to the market every day that makes a real-time normalized space domain implementation of correlation more feasible. Thus in this thesis a space variant approach to implement distortion invariant filters is discussed which highlights the main drawbacks present in the frequency domain approaches. Also effective optimization techniques are discussed to overcome the computational overhead associated with these approaches [11], [20].

One of the major drawbacks of frequency domain distortion invariant filters is that they start giving false detections in the presence of non-uniform light distributions and hence are not illumination invariant. Illumination invariance is a major requirement for all

pattern recognition systems which rely on real-time data acquisition and its processing for security and surveillance purposes.

This problem can be further illustrated as shown in figure 1.3 given below.

Figure 1.3 A non-uniformly illuminated scene

Global illumination normalisation would not work on this scene. If the target is present in the shaded areas of the scene, and structured clutter is present in the illuminated portions, than after correlation we may get a higher correlation peak for the structured clutter as compared to that of the real target.

Thus in Figure 1.3 it can be seen that because of the non-uniform light distribution, a filter to recognize the ball may fail, by producing high intensity false peaks in the region of high intensity clutter to the right of the image.

This can be demonstrated more quantitatively from the following simplified example shown in Figure 1.4 employing a frequency domain Optimal-Tradeoff Maximum Average Correlation Height (OT-MACH) filter trained on a non-uniformly illuminated

training dataset having a maximum pixel intensity of 40. In Figure 1.4 an autocorrelation function for the filer using the reference image having a maximum pixel intensity of 40 can be seen.

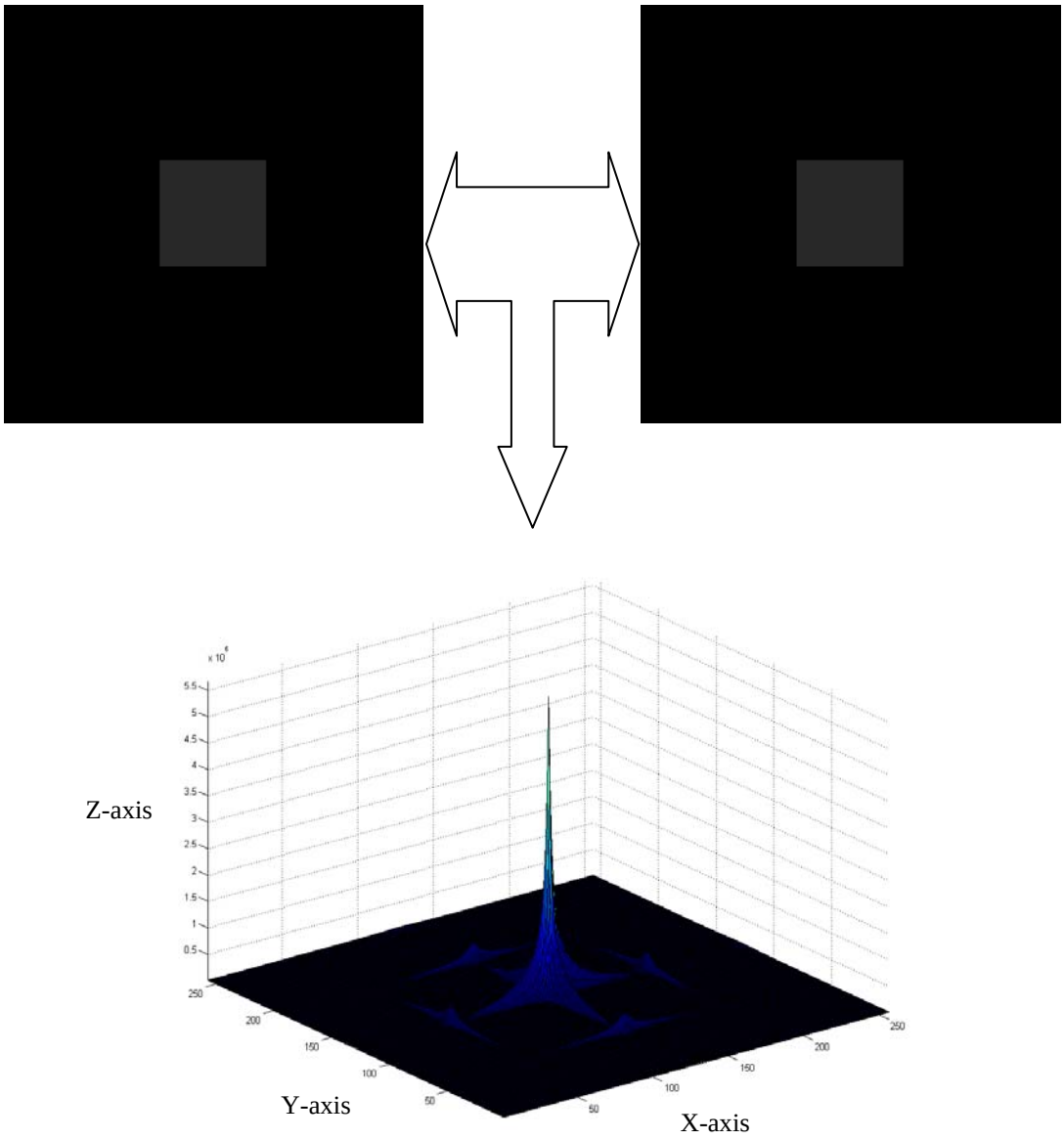


Figure 1.4 Autocorrelation using OT-MACH of training image dataset

The autocorrelation peak of the reference object yields a maximum peak of amplitude $5.65*10^5$ which can be seen from Figure 1.4 which shows in the correlation output plane produced.

The frequency domain OT-MACH is subsequently used to detect an illuminated false target having a maximum pixel intensity of 250.

In order to illustrate the limited capabilities of distortion invariant filters implemented in the frequency domain to be illumination invariant the following example is given. The OT-MACH filter was applied in the frequency domain for the correlation of the same reference image as before (i.e. a square of size 64×64 and pixel intensity 40) but with an illuminated target which was half the size and of increased pixel intensity of 250. The result is shown in Figure 1.5 below.

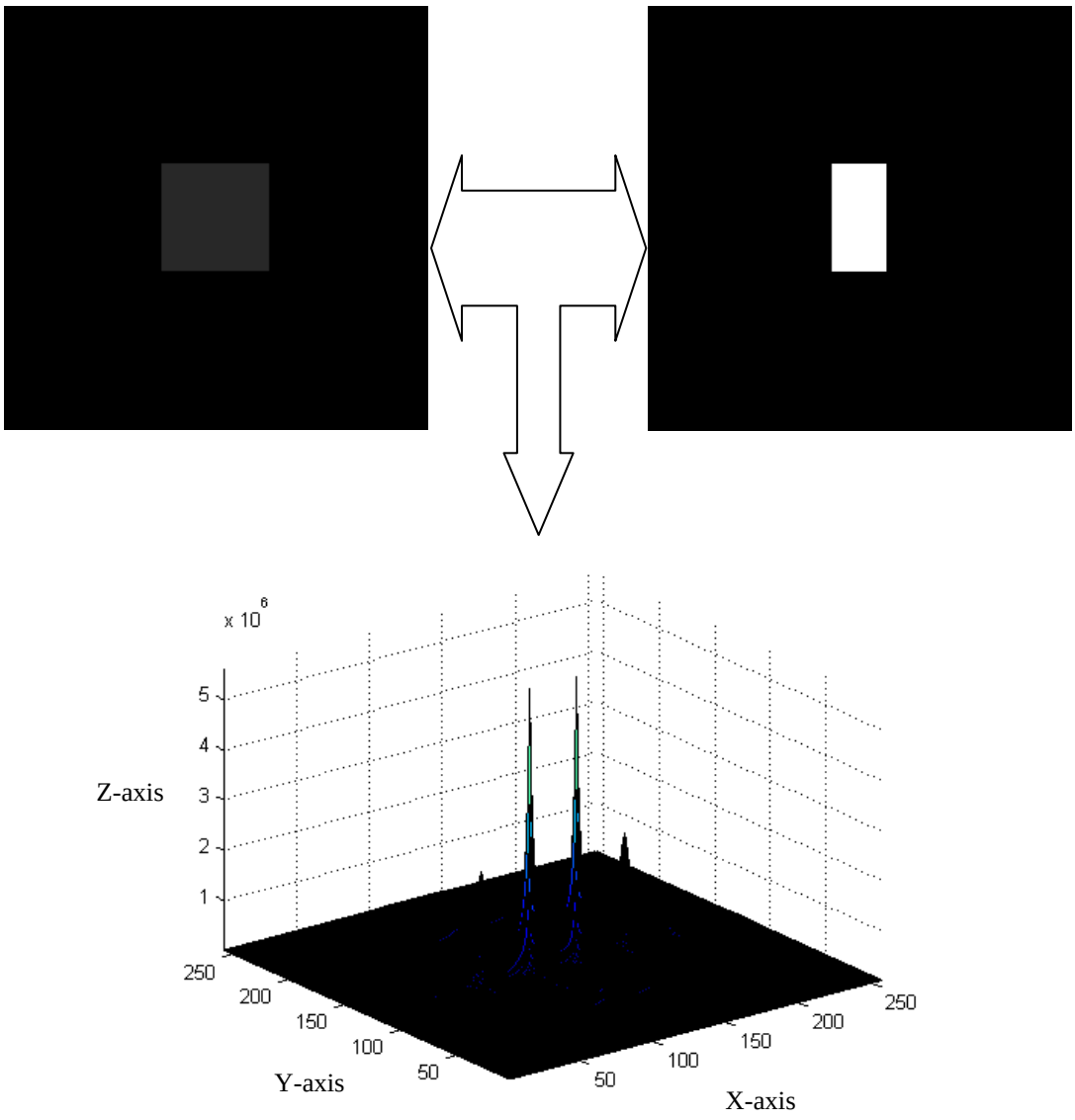


Figure 1.5 Correlation with OT-MACH using illuminated target

In this case not only are there dual peaks but also the amplitude of the highest correlation peak for the frequency domain OT-MACH filter implementation is $5.58*10^{6}$ which is more than the amplitude of autocorrelation function given by Figure 1.4.

Hence the frequency domain approaches fail in the scenario where illumination invariance is required. However this problem is not present in the spatial domain implementation where the local window energy normalisation creates a uniform distribution. This can be demonstrated in this example by using the equation 1.1 and 1.2 with the application of normalisation given by equations 1.6 and 1.7.

$$Normalised\_Correlation(T) = \frac{5.65\times10^{5}}{\sqrt{(64\times64\times40)^{2}\times(64\times64\times40)^{2}}} \tag{1.6}$$

$$Normalised\_Correlation(F) = \frac{5.58\times10^{6}}{\sqrt{(64\times64\times40)^{2}\times(64\times32\times250)^{2}}} \tag{1.7}$$

After the application of normalisation the value of $NORMALIZE(T) = 2.10*10^{-4}$ and for $NORMALIZE(F) = 6.56*10^{-6}$ which shows that the false target has a smaller correlation peak and hence the true class is that which is recognized. Although in this example the scenario considered in quite generic but in the forthcoming sections of this thesis a more detailed analysis will be performed considering different shapes and levels of shadings.

In this thesis a space variant OT-MACH filter is proposed employing local window energy normalisations along with optimization techniques to reduce the computational resources required. A comparison to existing techniques currently in the market is also made in this thesis.

1.4 Thesis Organisation

An effort has been made in this thesis to develop spatial domain filters which are invariant to the change in illumination. The only drawbacks associated with the spatial domain implementation techniques are their computationally intensive nature. The issues of computational intensity of the spatial domain filters are also addressed in this thesis by proposing speed enhancement techniques. A brief description of the thesis in terms of chapter break down is given below:

1. **Introduction**

   In this chapter the concept of correlation is given along with the drawbacks associated with frequency domain approaches. A problem area is defined and it is shown how illumination changes affect the performance of the frequency domain correlation filters.

2. **Distortion invariant correlation filters**

   In this chapter the history and development of correlation filters is discussed. An analysis is made of the digital implementation of correlation filters with reference to the previously used optical domain implementations. The idea of constrained and unconstrained correlation filters is discussed. A detailed analysis and results from a Synthetic Discriminant Function (SDF) filter applied to cluttered backgrounds are also given. Also in this chapter the need for unconstrained filters is made clear and a brief introduction to the Maximum Average Correlation Height (MACH) filter is also presented.

3. **Advancements in MACH filter**

   In this chapter the enhancements to the MACH filter considered in the previous chapter are discussed. Recent enhancements to the MACH filter which are known as an Extended MACH (EMACH) and Optimal Trade-Off Maximum Average Correlation Height (OT-MACH) filter which are aimed at the clutter rejection capability of the filter are also discussed.

4. **Spatial Domain Optimal Trade-Off Maximum Average Correlation Height (SPOT-MACH) filter**

   The space variant filter developed, which is named SPOT-MACH filter, is presented in this chapter. The capabilities of the SPOT-MACH filter in terms of its illumination invariance as compared to the frequency domain approaches are discussed in detail. The drawbacks of the SPOT-MACH filter in terms of computational intensity are also discussed in this chapter.

5. **Speed enhancement techniques for the SPOT-MACH filter**

   The SPOT-MACH filter is considered computationally intensive and in order to overcome this problem two algorithm based and two hardware based speed enhancement techniques are presented in this chapter. A timing analysis of the techniques is carried out with reference to their frequency domain counterparts to establish the effectiveness of the developed enhancements techniques.

6. **Applications of the SPOT-MACH filter and its comparison with Feature Based matching techniques**

   In this chapter the application of the SPOT-MACH filter to Forward Looking Infrared (FLIR) imagery and its hardware application for a security monitoring system is presented. Also, a detailed comparison of the SPOT-MACH filter is carried out with the popular feature based techniques. The dataset used for the techniques has been developed specifically to test the capabilities of the filters in response to varying illumination across the input scene.

7. **Conclusions and future work**

   In this chapter a detailed conclusion of the thesis and some proposals for future work in terms of hardware implementation of the SPOT-MACH filter and some further hardware based enhancement techniques are also summarised.

1.5 Conclusion

In this chapter, limitations are identified with the frequency domain implementation of correlation filters, their lack of tolerance to changes in illumination which can lead to false detections. An alternative approach to overcome this problem using a spatial domain implementation of the filter where the energy is locally normalised using a space variant moving window. The design and capabilities of the space variant filter based on this approach is discussed in detail in this thesis.

Chapter 2

Distortion Invariant Correlation Filters

2.1 Introduction

In the previous chapter the concept of correlation had been discussed and it was established what defines a good correlation. It was also described how correlation filters are an integral part of an IPR system. In this chapter different correlation filters are discussed in more detail to further elaborate on the defined problem area.

Digital computers or optical systems can be used for the implementation of IPR algorithms. Using an optical system for pattern recognition in the past has provided a temporal advantage over a digital implementation. Most of the digital pattern recognition operations are computationally intensive and rely heavily on the processing speed of the system. However, with the current advances in the digital domain the capabilities are quickly catching up with the high speed processing which has always been the major advantage of the optical domain. Another added advantage of implementation in the digital domain is the robust and generic nature of the developed algorithms with the options of rapid prototyping [9], [21].

On the other hand the optical techniques have always been known to provide parallelism, high speed processing, non-interfering communication and significant interconnection capability and so traditionally were considered a better alternative to digital implementation approaches [2].

2.2 History of Correlation Filters

Most of the initial correlation filters were developed using optical implementations. Optical pattern recognition techniques use correlation which offers unique advantages in pattern recognition applications owing to its sensor independent approach that does not require any of the data specific operations like segmentation, feature extraction and on-line model synthesis [9], [12].

Since the throughput of the correlator generally depends on the dimensions of the image, the algorithm complexity depends on the scene size and not the number of objects to be detected. This not only provides a precise estimate for the throughput, but it also avoids computational bottlenecks faced by some approaches due to complexity in the scene, clutter and the number of objects [9], [15] , [12] and [22].

It is known that correlation between two functions in the spatial domain is equivalent to conjugate multiplication in the spatial frequency domain. Thus in the implementation of an optical correlator the Fourier transform properties of a lens are used. An optical correlator normally contains two lenses to perform two Fourier transform operations successively. The first lens transforms the input image from the space domain to the spatial frequency domain and the second lens transforms it back to space domain. The input and the output functions are both related by a linear system in the space domain; this linear system can perform correlation or convolution. Apart from the lenses, other devices used in an optical correlator are Spatial Light Modulators (SLM), Charged Coupled Device (CCD) arrays, various mirrors and a beam splitter [3] , [5], [17] and [23].

When an optical correlator processes a target scene using the Fourier transform property of a lens, all the information of the input scene is compared with the reference image. The parallel processing of the optical correlator configuration provides a significant advantage in terms of processing speed over digital counterparts. The two most widely used optical correlators are namely the Joint Transform Correlator (JTC) and Matched Filter (MF) based Correlator [3] , [9] and [19].

In the case of a JTC an input image and a reference image are simultaneously Fourier transformed to yield a Joint Power Spectrum (JPS). The JPS is then inverse Fourier transformed to produce a correlation output. A match between the reference image and the input scene yields a pair of correlation peaks or bright spots at the output plane [9] , [24].

A classical JTC architecture is shown schematically in Figure 2.1 below:

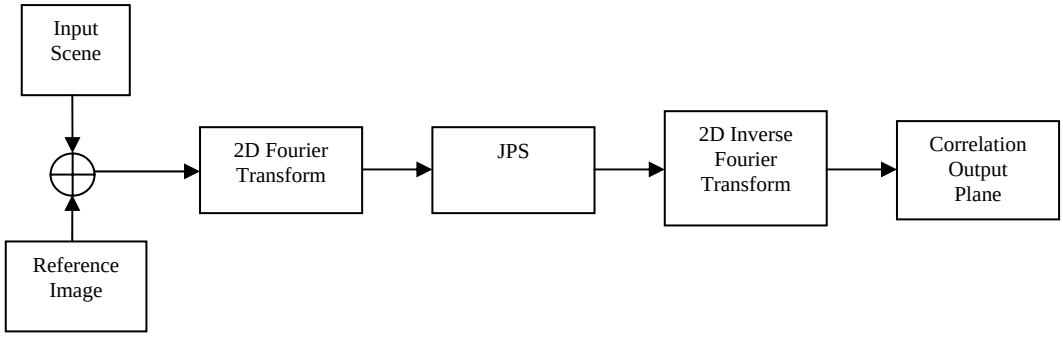


Figure 2.1 Architecture of Joint Transform Correlator (JTC)

The concept of the JTC emerged back in 1966, but the first real-time JTC implementation was introduced in 1984 with the development of real-time electronically addressable SLM. The classical JTC suffers from large correlation peaks width, large side lobes, strong zero order diffraction at the origin and false correlation peaks [17] , [25] and [26].

Over the years to enhance the performance of the JTC several improvements have been proposed such as the binary JTC, where the JPS is binarised based on a hard-clipping non-linearity at the Fourier plane to give only two values before applying the inverse Fourier transform to the correlation plane. Although it improves some aspects of the performance, but is also known to give slow processing speeds and harmonic correlation peaks [27] , [28].

To overcome the problems of the JTC some zero order diffraction elimination techniques such as Phase shifting method have also been used to eliminate the zero-order and to improve the poor detection capabilities [29] , [30].

Another notable enhancement to the JTC has been the use of Fourier plane JPS apodization based on the reference image. The resulting amplitude modulated JTS yields a better correlation output than its predecessors but also suffers from practical implementation difficulties especially if the power spectrum of the reference image contains zeros [31] , [32].

Alternatively, the matched filter is implemented in the frequency domain, the basic arrangement is shown in Figure 2.2 below:

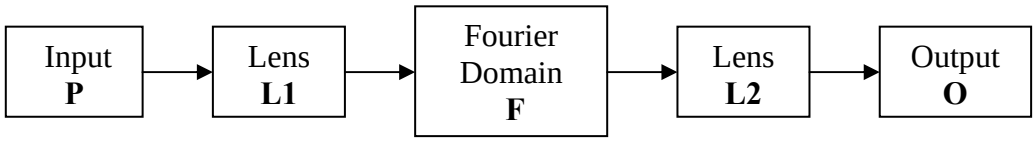


Figure 2.2 Matched Filter Schematic diagram

The input plane, P contains the input signal $f(\sigma,\varepsilon)$ and is illuminated by a monochromatic source. The plane P located at the front focal panel of the Fourier Transform lens $L1$. The complex light field distribution of the output plane F is given by [33]:

$$F(u,v) = \Im[f(x,y)] \tag{2.1}$$

where $\Im$ represents the Fourier Transform operation and σ and ε are spatial domain variables. Also $u = (2\pi/\lambda\mu)\sigma$ and $v = (2\pi/\lambda\mu)\varepsilon$ these are the frequency domain variables in which λ is the wavelength of the light source and μ being the focal length of the Fourier transform lens [3] , [5].

The phase fronts of the waves propagating from the plane P and incident on F tend to become a collimated wave front. The phase transmittance of the matched filter $H(u,v)$ at plane F must have a conjugated match to the phase of the Fourier Transform of $f(x,y)$. Thus the phase variations on the wave front incident on the plane F are

cancelled by the filter and all the resulting distributions beyond plane F becomes a parallel beam. The lens $L2$ performs a two dimensional Fourier Transform on the parallel beam and produces a correlation output on Plane O. In this case the correlation output is interpreted on the means of if the input signal is different from $f(\sigma,\varepsilon)$, than the phase front produced will not be completely cancelled by the filter. The light distribution at the plane O will effectively consist of a spot whose intensity represents the amplitude of the correlation between the two functions. The complex light distribution produced at the plane O has been given below [27] , [34], [33]:

$$g(x,y) = \Im'[*F(u,v).H(u,v)] \tag{2.2}$$

where $\Im'$ represents the two dimensional inverse Fourier transform operation on the wavefront. If the input object is moved across input plane, the Fourier transform remains fixed in space but is multiplied by a phase that depends on the lateral movement. The coordinates of the bright correlation spot in the output plane at O will then be proportional to the coordinates of the signal $f(\sigma,\varepsilon)$ location in the input plane [35], [36] and [36] .

The intensity of the bright correlation spot is proportional to the degree to which the input and the filter functions are matched which is also valid for multiple targets [8] , [33] and [37].

The matched filter, or classical Matched Filter (CMF), is thus a fully complex filter. To illustrate this, the frequency domain transfer function for a CMF is given by [5]:

$$H_{cmf} = R^*(u,v) = \left|R(u,v)\right| \exp\left[-j\Phi(u,v)\right] \tag{2.3}$$

where $\left|R(u,v)\right|$ is the amplitude and $\Phi(u,v)$ is the phase of the Fourier spectrum of the reference function $r(x,y)$. In the case when the input is similar to $r(x,y)$, the phase variation is cancelled in the Fourier plane and hence produces a collimated wave front. The correlation results from the CMF are not very sharp due to the squaring of the

magnitude $r(x, y)$. Also the filter is very sensitive to the small changes in the reference image [12] , [38] .

In order to further overcome the drawbacks of the CMF a new filter was proposed which is known as the Inverse Filter (IF), the frequency domain transfer function for which is given by [8] , [12] and [39]:

$$H_{if} = R(u,v)^{-1} = \left|R(u,v)\right|^{-1} \exp\left[-j\Phi(u,v)\right] \tag{2.4}$$

In the case of the inverse filter the correlation peak is sharpened since the Fourier amplitude becomes uniform after filtering. However, in the case where there are one or more poles the transfer function becomes undetermined which is considered a major drawback. In contrast, the transfer function of the Phase Only Filter (POF) in frequency domain is given as [36], [40] :

$$H_{pof}(u,v) = \frac{R^{*}(u,v)}{\left|R(u,v)\right|} = \exp\left[-j\Phi(u,v)\right] \tag{2.5}$$

The most optimum implementation of the matched filter in terms of light efficiency was the phase only filter (POF) which was designed by omitting the amplitude information [41] , [42].

After having had a brief look at the architectures and enhancements for matched filters and JTCs over the years it can be deduced that the transfer function of a matched filter and JTC enables detection in situations when an exact pattern needs to be matched in an input scene and will work provided only limited distortions are present. The detection capabilities of the matched and JTC filters are reduced progressively when the target object is rotated at angles away from the reference function. Thus the matched and JTC filters are known to fail in the presence of in-plane and out-of-plane rotations. Even when the object has not been rotated but variations in scale are present, the detection capabilities of the matched and JTC filter deteriorate [34] , [40] and [43]

Also the capabilities of these filters are limited in situations where multiple classes of target object are present in an input scene. In this scenario a single filter is unable to

detect the desired pattern and a filter bank might be needed which would be computationally intensive in nature [7] , [19] and [44].

Hence the need arose, for not only developing filters which could offer invariance to distortions in rotation, scale and the orientation of the target object but also allowed criteria to be defined for the quality of detection. [45] , [46].

Advances in the field of optical pattern recognition led to the design of a new class of filters providing tolerance to in-class variations and discrimination between members of different classes [16], [22], [28] and [47].

2.3 Performance Metrics for correlation filters

In order to define the criteria of detection it is necessary to calculate some basic measures to represent the quality of the output correlation plane.

The most basic measure of the correlation plane is the correlation output peak intensity (COPI) which is defined as [39]:

$$COPI = \max\left\{\left|C(x,y)\right|^2\right\} \tag{2.6}$$

where the value of C(x, y) is the amplitude of output correlation plane at position (x, y).

To provide the best detection capability and performance it is necessary for a filter to yield a sharp correlation peak as well as a high COPI value, while keeping side lobes to a minimum. The filter's ability to do this can be measured using the peak-to-correlation energy (PCE) measure [39] , [48].

The basis of the PCE measure is that the correlation peak intensity should be as high as possible while the overall correlation energy in the plane should be as low as possible. A high PCE value therefore implies that the filter performs well. The PCE ratio is generally defined as:

$$PCE = \frac{COPI}{Energy} \tag{2.7}$$

where *Energy* is the total correlation plane energy and is defined as:

$$Energy = \sum |C(x, y)|^2 \tag{2.8}$$

Also another parameter in defining the performance of a correlation filter is Peak to Side Lobe Ratio (PSR) which forms an important criterion especially in the case when multiple objects are present in the input scene. In most cases considering just the COPI values would not be enough to analyze the performance of the filter [39]. The PSR is calculated by subtracting the mean of the correlation plane from the COPI and than dividing it by the standard deviation of the correlation plane.

$$PSR = (COPI - MEAN) / \sigma \tag{2.9}$$

where σ denotes the standard deviation of the correlation plane intensity values. The PSR of the target object should be greater than background clutter. It can be used in the case when the difference between the two COPI values of the target and reference objects is negligible [49] , [50].

PSR is a metric that has been widely used to measure the sharpness of the correlation peaks. When the target belongs to a true class the PSR values should be high and in the case of false class objects the PSR value should be low. Now in order to effectively compute PSR the test image is cross-correlated with the reference image and the output correlation plane is scanned for the largest value. When computing the detection criteria many projections should be used instead of a single projection of inner products as this enables the correlation filter to produce a specific response facilitating an accurate detection. It has also been observed that the PSR metric is invariant to changes in illumination of the training images [51] , [52] and [53].

2.4 Correlation filters in frequency domain

A desired target in an input scene can be distorted in many different ways including in-plane or out-of-plane rotation, variation in scales, background clutter, noise, low illumination or in the case of outdoor video sequences constantly changing lighting angles or shadows [54] , [55].

There are many different ways to achieve distortion tolerant pattern recognition for use with a matched filter based Correlator. Capabilities and limitations of some of the most widely used methods are discussed below.

2.5 Synthetic Discriminant Function (SDF) Theory

The SDF theory has been one the most influential in the field of correlation pattern recognition for many years. One of the earliest composite correlation filters was actually known as the SDF filter [8].

SDF theory has been developed on the principle that the expected distortions the targets are included in the design of the filter [45].

Thus in the SDF theory the filter is designed to yield a specific value at the origin of the correlation plane in response to each training image. In a two-class problem, the correlation values at the origin can be set to 1 for training images from one class and to 0 for the training images from the other class [56]. The aim is that the filter will yield values close to 1 for all images from class 1 and close to 0 for all images from class 2. So in the case of detection the class of the target can be deduced by the value at the origin of the correlation plane [5].

This approach will provide a certain level of invariance to the expected distortions. Another main advantage of the SDF is that objects from multiple classes can be included in the filter design to make it more robust in scenarios where more than one object is to be detected [8].

The conventional SDF filter is designed on the principle that the weighted linear combination of reference images is used to create a linearly superimposed composite image [35] , [57]. The composite image has to be cross correlated with the input image containing the target object in the presence of background clutter. This will give the output correlation plane containing on-axis equal correlation peaks for all target objects belonging to the same class. A sample dataset used in this thesis containing a single class of object, i.e. a car, is shown in Figure 2.3.

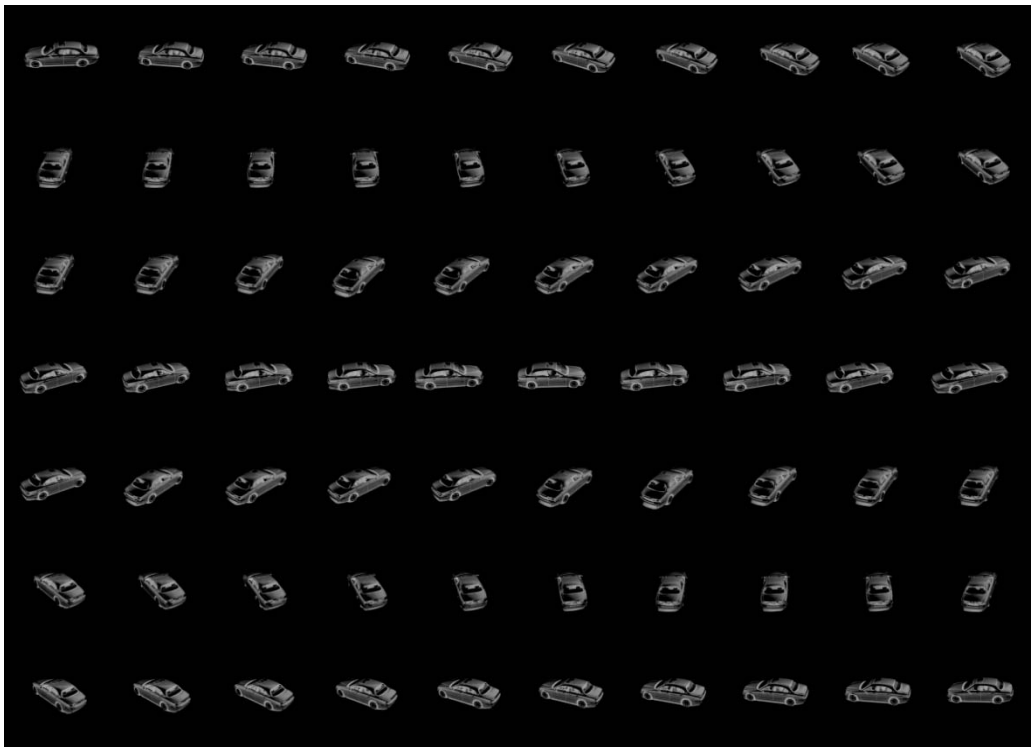


Figure 2.3 Training image data set

Now in this case $t_i(x,y)$ denotes the training image data set as shown in Figure 2.3 where $i = 1,2,\ldots\ldots,N$ and N is the number of the training images to be used in creating a composite image which can be denoted as $h(x,y)$. The values at the origin of the output correlation plane are set equal to a constant. The output correlation plane between the composite image and the training images has been defined in equation 2.10 given below [45] , [58] and [59]:

$$h(x,y).t_i(x,y) = \int\int t_i(x,y)h^*(x,y)dxdy = c \tag{2.10}$$

Now as the composite image is defined as a linear combination of the N training images, as shown in equation 2.11 [45], [60]:

$$h(x,y) = a_1 t_1(x,y) + a_2 t_2(x,y) + ... + a_N t_N(x,y) \tag{2.11}$$

where a_N represents the weighted mean of N training images.

By further simplification, equation 2.11 can be written as:

$$h(x,y) = \sum_{i=1}^{N} a_i t_i(x,y) \tag{2.12}$$

By substituting equation 2.12 in 2.11:

$$\sum_{i=j=1}^{N} a_i^* R_{ij} = c \tag{2.13}$$

where

$$R_{ij} = \int\int t_i(x,y)t_j^*(x,y)dxdy \tag{2.14}$$

In this case R_{ij} is defined as the correlation output at the point of origin for the training image date set. In the situation where multiple classes of objects are present, containing actual and false targets, the filter is required to accept the true targets and reject the false ones. The training images for the false targets can be included in the design of the filter such that the cross correlation value between the true and false class training images is close to zero. A composite image created from the training image dataset in Figure 2.3 is shown in Figure 2.4:

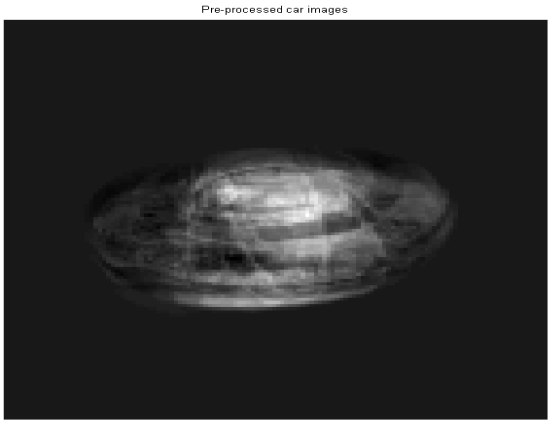


Figure 2.4 $0^{\circ} - 360^{\circ}$ orientations over 10° interval SDF composite Image

In order to determine the impulse response of the filter, as shown in equation 2.15, the composite image has been correlated with a car image orientated at 0 degrees without any background. The result is shown in Figure 2.5 below:

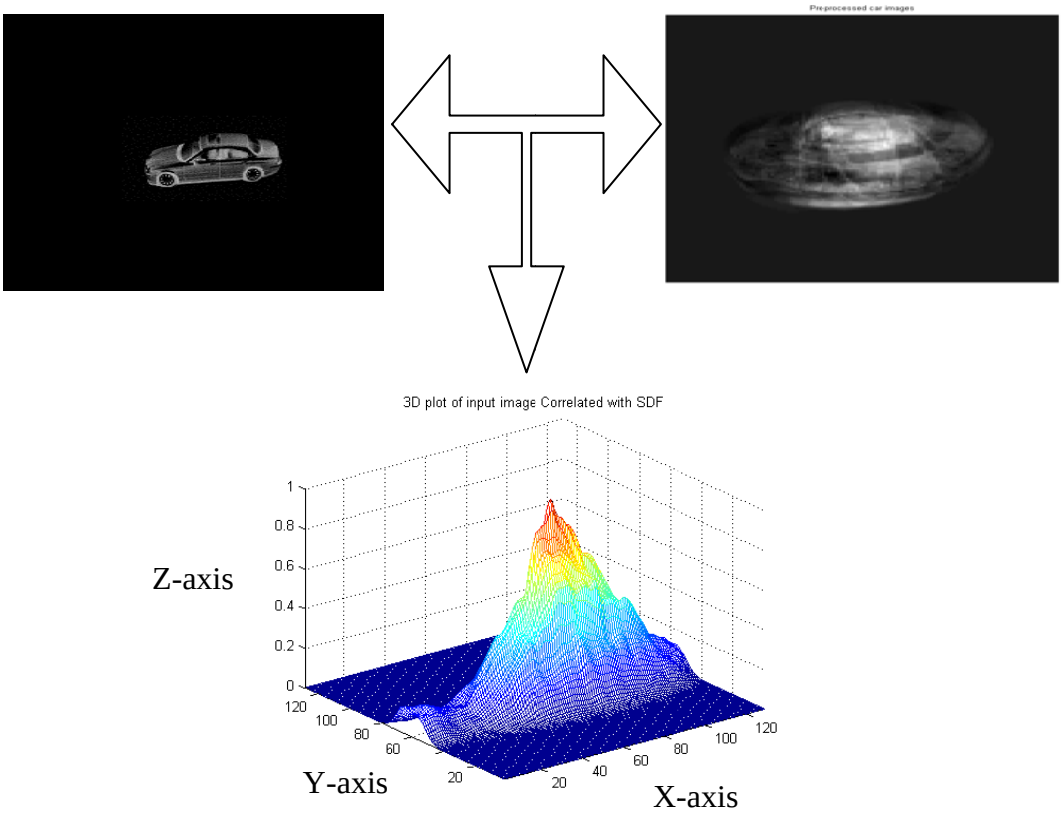


Figure 2.5 SDF correlation output

The performance ratios for the correlation output plane given in Figure 2.3 are given in Table 2.1:

PCE	PSR	Orientations (Target , Reference)
0.0976	1.1269	(0,0-360) + No Background Clutter

Table 2.1 Performance ratios for SDF Correlation Plane

From Table 2.1 it can be seen that the correlation output plane for the SDF has a very low PSR value showing that there is a very broad peak and the presence of large side lobes.

In order to evaluate the discrimination ability in the scenario where the target is oriented at an angle not present in the design of the filter a target object has been oriented at an arbitrary angle of 5 degrees in the absence of any background clutter. The correlation output plane is shown in Figure 2.6:

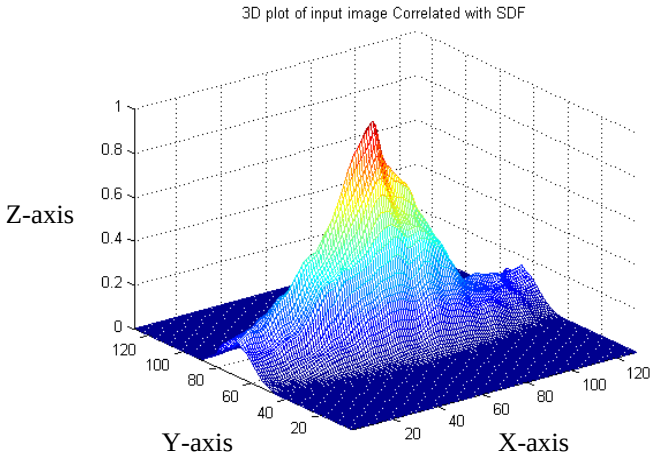


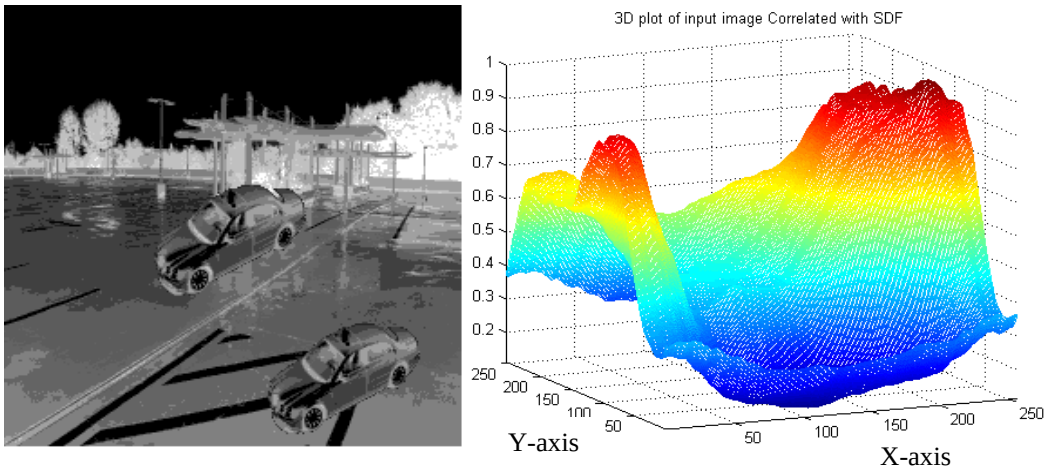
Figure 2.6 Correlation output plane for orientation invariance

The performance ratios for the correlation plane shown in Figure 2.4 are given in Table 2.2:

PCE	PSR	Orientations (Target , Reference)
0.0939	1.0601	(5,0-360) + No Background Clutter

Table 2.2 Performance ratios for SDF Correlation Plane for distortion invariance

Although the SDF filter design is known to provide invariance to distortions, it is not optimal in situations where background clutter is present [18], an example response to a cluttered scene being shown in Figure 2.7:



(a) **(b)**

Figure 2.7 Multiple targets in clutter background (a), Correlation plane (b)

From Figure 2.7 it can be seen that in the presence of background clutter the discrimination abilities of the SDF fail.

Thus a conventional SDF can give good distortion tolerance but from Figure 2.7 it can be seen that it is not optimal for clutter tolerance. Also it produces broad peaks in the output correlation plane. Different modifications have been proposed in SDF design to improve upon this performance some of which are discussed in the next section.

2.6 Advancements in SDF theory

Most of the SDF methods have been generalized in both frequency and spatial domains [45], [61].

A 'd' dimensional column vector $s_1, s_2, ..., s_N$ is acquired by lexicographical scanning. The d dimensional column vector is denoted h and used to describe the composite image $h(x, y)$; then the Equal Correlation Peak (ECP) constraint can be written as:

$$h^{\tau} S_i = c_i \tag{2.15}$$

where τ denotes the conjugate transpose and S_i is the column vector containing $s_1 - s_N$.

$$S^{\tau} h = c^* \tag{2.16}$$

The composite image *h* from the SDF is of the following form:

$$h = S.a \tag{2.17}$$

Where '*a*' is the vector of coefficients defined before. Using the above two equations and solving for '*a*' yields:

$$a = (S^{\tau} S)^{-1} c^* \tag{2.18}$$

$$h_{ECP} = S(S^{\tau} S)^{-1} c* \tag{2.19}$$

By using the generalized inverse method that gives a general expression for any *h*, equation 2.19 can be written as:

$$h = x_p - x_{h0} \tag{2.20}$$

where $x_p = S(S^{\tau} S)^{-1} c^*$ and $x_{h0} = [I_d - S(S^{\tau} S)^{-1} S^{\tau}] z$ where *z* is any 'd x *1*' vector and I_d is the *d* x *d* identity matrix [45].

The first breakthrough in the rigorous development of SDF theory came with the development of the Minimum Variance Synthetic Discriminant function (MVSDF) [62] , [63] .

The MVSDF improves the noise tolerance of the SDF by using the knowledge of the noise covariance matrix (X). This technique minimizes the variance of the correlation output and maximizes the signal to noise ratio. The MVSDF is defined by the following equation [62] , [64]:

$$h = X^{-1}T(TX^{-1}T)^{-1}c \tag{2.21}$$

where T is a matrix where each column represents a single training image after lexicographical scanning. To improve the accuracy of the filter the noise covariance matrix must be known. Also the MVSDF cannot control the side lobes in the correlation plane and therefore a false detection may result. The MVSDF was found to be useful for minimising the effects of additive input noise. Nevertheless the use of the MVSDF was considered impractical owing to the need for inverting a large covariance matrix; However the optimization approach opened the doorway for the advancement of SDF theory [8], [60].

Following the advancements of the MVSDF, Mahalanobis *et al.* proposed the Minimum Average Correlation Energy (MACE) filter [62], which produced sharp peaks for easy detection of the filter output. The MACE filter is generally known to be sensitive to distortions but readily able to suppress clutter [65].

MACE filters are synthesised efficiently in the frequency domain with the help of closed form equations. These equations can be obtained by using a criterion that seeks to minimize the average correlation energy computed from cross-correlation of the training image data-set while satisfying certain linear constraints to produce a specific value at the origin of the correlation plane for the given training images. The MACE filter tries to control the entire correlation plane instead of a single point. It minimizes the energy of the entire plane to produce sharp peaks at the origin and good side lobe

suppression by producing values close to zero where the co-ordinates of the peak indicate the location of the target. In general terms the MACE filter can be defined as a two stage filter where the first stage uses a transfer function related directly to the average power spectrum of the training images and the second stage is a SDF filter based on the training data-set filtered in the first stage. The optimal MACE filter is given by the closed form equation 2.23 given below [18] , [65] and [66]:

$$h = D^{-1}T(T^{+}D^{-1}T)^{-1}c \tag{2.22}$$

Here 'D' is the diagonal matrix with the assumption that none of the values of the diagonal correspond to zero. So it can be assumed that 'D^{-1}' is a diagonal matrix as well. Since the diagonal elements of 'D' are equal to the average power spectrum of the training images so the elements of 'D^{-1}' are the reciprocals of the average power spectrum [65].

Now it can be assumed that $D^{-0.5} = P$, where P is a diagonal matrix with its diagonal elements being reciprocal of the square roots of the diagonal elements of D. This can be seen in equation 2.23 given below:

$$h = P(P.T)(T^{+}.P.P.T)^{-1}c \tag{2.23}$$

Further it can be assumed that $PT = \bar{T}$, so substituting in equation 2.23 gives:

$$h = P\bar{T}(\bar{T}^{+}\bar{T})^{-1}c \tag{2.24}$$

Now equation 2.25 can be further simplified when $\bar{h} = \bar{T}(\bar{T}^{+}\bar{T})^{-1}c$ which is shown in equation 2.25 given below:

$$h = P\bar{h} \tag{2.25}$$

From equation 2.25 it can be seen that $\bar{h}$is a SDF based on the transformation training data $\bar{T}$. So in the frequency domain h can be represented in cascaded form of P and the projection of SDF $\bar{h}$.

In short the MACE filter can be described as being an ordinary SDF filter operating on pre-processed data. In this case the pre-processing stage is makes the average power spectrum of the entire training images equivalent to white noise.

Thus the MACE filter is considered an advancement of a conventional SDF using transformed training data making it suitable for environments with background noise where P is acts as a whitening filter at the pre-processing stage.

But in the case of the MACE filter sharpness of the peak improves the discrimination ability of the MACE filter but produces poor performance in the case of distorted targets that have not been included in the design of the filter [51] , [65].

As observed from the MACE filter the detection capabilities of the SDF filter depends largely on the training set used. For optimal detection the training set must contain all the expected distortions. But the problem is that the value of N should be small so that the matrix size is small. It has been observed that it is not necessarily beneficial to keep on increasing the value of N. As after a certain value of N the performance of the SDF begins to degrade [67].

The main factor in SDF theory is that the distortion tolerance can be controlled by using the desired correlation peak values.

The values can be taken as 1 for the in-class objects and 0 for the out-of-class objects. The introduction of the non-target objects in the composite image tends to improve the class detection ability of the SDF filter.

Apart from SDF theory there are alternative approaches which tend to remove the constraints in the SDF to enable more robust filter design, resulting in the generation of unconstrained correlation composite filters.

2.7 Unconstrained Correlation Filters

The SDF filters discussed initially have been able to exhibit advantages like sharp correlation peaks or tolerance to background noise. But the SDF filters were unable to demonstrate optimization capabilities to offer good distortion tolerance alone. In this case the choice of training set greatly influences the spectral abilities of the filter which influences its ability to tolerate distortions. It has also been observed that filters which produce broad correlation peaks are able to offer better distortion tolerance but reduced performance in the case of discriminating against classes. The reason for this might be that these filters offer correlation with low-frequency information from which classes might not be easily discriminated. Also another method to increase distortion tolerance of the filter is to increase the number of training images but this makes the filter more computationally intensive and in many cases acquiring a large data set of training images is problematic [9] , [68].

The SDF theory has been based on the assumption that the distortion tolerance of a filter could be controlled by explicitly specifying desired correlation peak values for training images. Thus in SDF theory, as has been mentioned above the correlation peak values are assigned as 1 for true class and 0 for false class objects. The SDF approach is constrained in this regard as the values of the peaks are restricted [5] , [13] and [38].

Now another method is to remove these constraints all together as non-training images always produce different results from those specified for the training images.
Also there is no relationship between the constraints and the ability of the filter to be invariant to distortions. In this case the correlation plane can be treated as a new pattern generated from the filter as a response to the input image. Hence the correlation planes are given us the linearly transformed versions of the input image. And attention has been paid not only to the correlation peak but the entire area of the correlation output plane [9] , [65].

The extension of the idea of the SDF led to the development of the Maximum Average Correlation Height (MACH) filter. The detailed design and capabilities of the MACH filter and its enhancements are discussed in more detail in forthcoming chapters. In brief the MACH filter finds the optimal compromise between the pattern discrimination

criteria in the presence of background noise. The MACH filter maximises the relative height of the average correlation peak with respect to the expected distortions included in the design of the filter [9] .

The MACH filter achieves distortion tolerance by maximising the similarity in shape of true class correlation outputs. This is done by the minimisation of the metric known as the average similarity measure (ASM). MACH filter designs also maximise the average correlation peak height (ACH) for in-class objects and minimise the average correlation energy (ACE) for out-of-class objects. These filters depend on the average training image, which may resemble clutter more than a target and lead to poor clutter rejection [9], [46] .

Unlike the SDF, the MACH filter can be tuned using the weighted coefficients to maximize the correlation peak height, peak sharpness and noise suppression whilst also being tolerant to distortions in the target image [9], [68].

2.8 Conclusion

In this chapter distortion invariant correlation filters have been discussed. A brief overview of the SDF filter and its capabilities has been given in the frequency domain. It has been discussed how the MACE filter was the first SDF type filter to provide control over the shape of the output correlation plane by minimizing correlation energy using linear constraints. The drawbacks in SDF theory have been discussed with regard to the lack of provision of distortion tolerance and hence invariance to the change in orientations of the target object.

Also in this chapter a new type of unconstrained filters has been discussed which generalizes the concept of ACE using a metric called ASM. The ASM criterion was considered as one of the first approaches to enable filters to be invariant to distortions. The enhancements of this approach has led to the creation of the MACH filter.

The MACH filter, which has been briefly discussed in this chapter, has the ability to minimize distortions using a mean-squared-error criterion. A lot of variations have been made on the MACH filter over the years which have resulted in performance

improvements, some of which will be discussed in the forthcoming chapters along with our own proposed techniques.

In conclusion, the correlation filter design has been shown to offer better invariance to distortions by the use of relaxed constraints as compared to hard constraints. There is always a compromise in the design of the correlation filters in terms of making them robust to varying noise criteria, resulting from the presence of unknown clutter. Hence the introduction of unconstrained distortion invariant filters has been made.

Chapter 3

Advancements in MACH Filter

3.1 Introduction

In the previous chapter, an overview of correlation filters was discussed along with their optical implementations. The limitations of the classical matched filter and the JTC were also discussed along with a brief overview of enhancements made over the years.

Also the constrained correlation criterion was discussed and a digital implementation of a SDF filter along with its performance evaluation in a cluttered background was also discussed.

In the end it was concluded that certain limitations in the constrained correlation criteria makes it a non viable option in the presence of background clutter. The unconstrained correlation criteria was also discussed and its advantages. The MACH filter was introduced as a distortion invariant correlation filter which employs unconstrained correlation criteria.

In this chapter the digital implementation of the MACH filter in the frequency domain is discussed along with its enhancements and limitations.

3.2 Unconstrained Correlation Criteria

As has been discussed in the previous chapter the SDF filters performance is controlled by imposing hard constraints on the output of the filter. The method for selecting constraint values for the design of the SDF filter has been largely unaddressed. The MACH filter avoids this problem by acquiring a large average peak without assigning constraints for individual training images employing unconstrained correlation criteria.

To understand the design of a MACH filter consider a training set which consists of N images of size $p1 \times p2$ where each image contains $p = p1 \cdot p2$ pixels. The i^{th} training image for the target class is denoted by $x_i(m,n)$ and the frequency domain representation is given by a $p \times 1$ vector X_i where here X_i can be obtained by lexicographically arranging its two dimensional Fourier transform given by $X_i(k,l)$. The Fourier domain filter is given as a $p \times 1$ vector h and its two dimensional representation is obtained by rearranging h in the form of an image which is denoted by $H(k,l)$.

The frequency domain correlation of the i^{th} training image and the filter has been given as [69]:

$$g_i = X_i \cdot h \tag{3.1}$$

In this case X_i is a $p \times p$ diagonal matrix containing the elements from vector x_i. The deviation in the shape of the correlation plane with respect to an ideal shape vector u and is quantified by the Mean Squared Error (MSE) given by equation 3.2 below [9] , [64]:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(g_i - u)^{+}(g_i - u) \tag{3.2}$$

where MSE is the measure of distortion with respect to the reference shape vector u which can be chosen as desired. The shape vector u plays a very important role in the distortion minimization problem as it can also be applied in the design of the SDF filter which has been discussed in the previous chapter to improve its performance. This is known as a MSE-SDF filter and can be achieved by using the shape vector u in the form of a Gaussian ring in order to re-shape the correlation plane which in turn effectively suppresses the noise in the correlation plane by only concentrating on a desired region of interest [64].

In the design of the MACH filter the shape vector u is chosen such that it causes the least amount of variation within the correlation planes and offers minimum MSE. In

order to find the most optimal shape vector u_{opt} the gradient of MSE with respect to u is set to zero which can be seen in equation 3.3 given below [7] , [9]:

$$\nabla_u(MSE) = \frac{2}{N}\sum_{i=1}^{N}(g_i - u) = 0 \tag{3.3}$$

From equation 3.3 u_{opt} can be given as [9]:

$$u_{opt} = \frac{1}{N}\sum_{i=1}^{N} g_i = \bar{g} \tag{3.4}$$

where $\bar{g}$ can be denoted as [9]:

$$\bar{g} = \frac{1}{N}\sum_{i=1}^{N} X_i \cdot h = M \cdot h \tag{3.5}$$

And where $\bar{g}$ is the average correlation plane and M is the average training image given as a diagonal matrix which can be elaborated by equation 3.6 [9]:

$$M = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{3.6}$$

So it can be deduced that $\bar{g}$ offers the smallest possible MSE and the least distortion within the correlation planes. In order to yield better performance $u$can be substituted with $\bar{g}$ in equation 3.2 which gives the expression for the Average Similarity Measure (ASM) given below [9]:

$$ASM = \frac{1}{N}\sum_{i=1}^{N}(g_i - \bar{g})^{+} \cdot (g_i - \bar{g}) \tag{3.7}$$

which can be further elaborated as [9]:

$$ASM = \frac{1}{N}\sum_{i=1}^{N}(X_i h - Mh)^{+} \cdot (X_i h - Mh) \tag{3.8}$$

Further factorization with respect to h yields [9]:

$$ASM = h^{+}\left[\sum_{i=1}^{N}(X_i - M)^{+}\cdot(X_i - M)\right]h = h^{+}S_x h \qquad (3.9)$$

where h represents the correlation filter, S_x being a diagonal matrix measuring the similarity of the training images to the class mean in the frequency domain. In the case all training images are identical then S_x would be a zero matrix. Also from Parseval's theorem it can be shown that the average squared distance from the correlation planes to their mean is the same as given by equation 3.9 in the frequency domain. Similarity matrix S_x can be further elaborated by equation 3.10 [67] , [70] and [71]:

$$S_x = \sum_{i=1}^{N}(X_i - M)^{*}(X_i - M) \qquad (3.10)$$

The nature of most unconstrained correlation filters can be loosely classified as to define the transfer function h that tends to minimize the Average Similarity Measure (ASM) across all possible circular shifts of the training image data set subject to some constraint at the origin of the data set [9] , [71].

The ASM has been proven to measure the effect of different types of distortion on the performance of the filter. For an ideal filter the ASM must be minimized and for this MSE should be taken into consideration. Now from the above equations it can be deduced that ASM is the measure of distortions in the correlation plane relative to an average shape. Ideally all the correlation planes should be identical for distortion invariant unconstrained correlation filters, and the value of ASM should be zero. Minimizing the value of ASM improves the stability of the filter's output response in the presence of distortions [5] , [72] and [73].

As discussed above ASM is one possible criterion for distortion tolerance as it represents the average deviation of the correlation planes from the mean correlation shape given by $\bar{g}$. This also gives the compactness of the class in the case filter $h$is viewed as linear transform then ASM would measure the distance of the training images

from the class center under the same transform. Minimization of *ASM* leads to a compact set of correlation planes which exhibit the least possible variations. The distortions of the reference object in the input plane are represented by the set of training images x_i. These distortions are reflected in the output as variations in the structure and shape of the corresponding correlation planes g_i and are quantified by *ASM*. In the case the filter reduces the distortion then the cluttered input image should yield similar output planes leading to a small value of *ASM* [9] , [68].

The MACH filter in principle relaxes the correlation peak height constraints and maximizes the peak intensity of the average training image. This has been given as another important measure in the design of unconstrained correlation filters known as the Average Correlation Peak Height (ACH). This has been achieved by maximizing the correlation height by correlating with the average of the training images and maximizing the magnitude of the resultant peak. The peak intensity of the average training image is given by $|\bar{g}(0,0)|^2$ which has been further elaborated by equation 3.11 [9].

$$ACH = |\bar{g}(0,0)|^2 = \left| \frac{1}{N} \sum_{i=1}^{N} h^+ \cdot x_i \right|^2 = \left| h^+ \cdot m \right|^2 = h^+ m m^+ h \qquad (3.11)$$

where m is the Fourier transform of the average training image in the vector form, also it has been assumed without the loss of generality that the peak occurs at the origin of the correlation plane. As discussed above the smaller the value of *ASM* the more invariant the response of the filter. So it can be deduced that if *ASM* is small then all true-class correlation planes would be similar to $\bar{g}$. So it is required by the filter h to produce a sharp correlation peak with the mean image while also minimizing the value of *ASM*. But for a robust filter it must possess the ability to reduce the noise variance at the output to some degree. The Output Noise Variance (*ONV*) for filter h is given by [9]:

$$ONV = h^+ D h \qquad (3.12)$$

The ONV is given by equation 3.12 where D is the diagonal power spectral density, although practical noise may be more multiplicative and tends to make the background more cluttered as compared to the additive noise but the additive noise model provides some level of robustness. So from the performance measures discussed above the MACH filter can be denoted by equation 3.13 [9] , [68] :

$$J(h) = \frac{(ACH)^2}{ASM + ONV} \tag{3.13}$$

Substituting the value of ACH in equation 3.13 [9]:

$$J(h) = \frac{\left|\bar{g}(0,0)\right|^2}{ASM + ONV} \tag{3.14}$$

Further simplifying equation 3.14 by substituting the values [9]:

$$J(h) = \frac{\left|h^+ \cdot m\right|^2}{h^+ \cdot S \cdot h + h^+ \cdot D \cdot h} \tag{3.15}$$

or,

$$J(h) = \frac{h^+ \cdot m \cdot m^+ \cdot h}{h^+ \cdot (S + D) \cdot h} \tag{3.16}$$

The optimal transfer function for $J(h)$ can be achieved by finding the derivative of equation 3.16 with respect to h and setting to zero [9].

$$h = (S + D)^{-1} \cdot m \tag{3.17}$$

The transfer function h given by equation 3.17 is known as the MACH filter transfer function as it maximizes the height of the mean correlation peak relative to expected distortions. In the case where an estimate of D is not available a white noise covariance

matrix is used by substituting $D = \sigma^2 \cdot I$ where I being the diagonal identity matrix. So by equating the value of D is equation 3.17 the new simplified MACH transfer function can be given by equation 3.18 [9]:

$$h = (S + \sigma^2 \cdot I)^{-1} \cdot m \tag{3.18}$$

By substituting σ^2 with a constant γ:

$$h = (S + \gamma \cdot I)^{-1} \cdot m \tag{3.19}$$

The MACH transfer function in its simplest form can be seen from equation 3.19, it can be deduced that the robustness of the MACH filter can be attributed to the inclusion of the *ASM* criterion in the denominator of the transfer function along with *ONV* which reduces the sensitivity to distortions and contributes to the removal of the hard constraints on the correlation peak. The MACH filter transfer function has the capability to adjust the correlation planes to a suitable value for optimizing the performance. The MACH filter's ability to handle distortions depends largely on the expected distortion within the training set which are used in the design of the filter.

In the case of an IPR system the MACH filter given in equation 3.19 can be converted to a two dimensional form which can be expressed as $H(u,v)$. Suppose there is a two dimensional target image given by $z(m,n)$ than it should be Fourier transformed to obtain $Z(u,v)$. For an IPR system the frequency response $Z(u,v)$ of the target image should be then multiplied with the two dimensional filter given by $H(u,v)$ in the Fourier domain which can be seen from equation 3.20:

$$G(u,v) = Z(u,v) \cdot H^*(u,v) \tag{3.20}.$$

where $G(u,v)$ is the product of the two spectra. The spatial domain correlation output can be achieved by applying the inverse Fourier transform to the equation 3.20.

$$g(m,n) = \left| \Im^{-1}\left[G(u,v) \right] \right|^2 \tag{3.21}$$

The spatial output of the correlation plane can be seen from equation 3.21 which records the intensity values from which the location of the target object can be deduced. The MACH filter discussed above utilizes the unconstrained correlation criteria to detect target objects from cluttered backgrounds. Although robust in nature the classical MACH has room for improvements and hence over the years a lot of advancements have been proposed. In this section the design and understanding of a simple MACH filter has been discussed, further advancements to the MACH filter are discussed in the forthcoming sections of this chapter.

3.3 Enhancements to the MACH filter

It has been discussed above that the natural extension of the MSE-SDF led to the basis of formulating a MACH filter, which determines and uses the correlation shape yielding the minimum squared error. In order to improve the clutter rejection capabilities of the MACH filter a lot of modifications have been proposed over the years.

It has been discussed earlier that the MACH filter is designed to maximize the intensity of the average correlation output at the origin. This leads to the fact that the average correlation peak is the output due to the average training image. Also in the case of the MACH filter it also maximizes the similarity between the average training image's correlation output and the output due to all training images from the desired class. It is evident from this that the MACH filter tends to force the images from the desired class to follow the behavior of the average training image and hence relies heavily on the mean training image. From this it can be deduced that the MACH filter enhances lower frequency components while suppressing higher frequency within the training set. So ideally a MACH filter can be obtained by using the mean image m as the only sample that represents all the training images. But the filter that has been obtained from this process may fail to capture the finer details of the training images and as a result may fail to discriminate the desired class from cluttered backgrounds. The reliance of the MACH filter on the average training image in some cases would lead to the detection of clutter as the desired target. This is due to the fact that using the average training image as the only criterion of detection is not a good practice as the m might not necessarily be a good representative of the desired class [74].

A more recent enhancement to the MACH filter to overcome this problem is known as an Extended MACH (EMACH) which is aimed at the clutter rejection capability of the filter [9]. In this case to control the relative contribution of the desired class training images and the mean image a new metric has been introduced which is known as All Image Correlation Height (*AICH*) which can be defined by equation 3.22 below [74].

$$AICH = \frac{1}{N}\sum_{i=1}^{N}\left[h^{+}(x_i - \beta m)\right]^2 \tag{3.22}$$

From equation 3.22 β is constant and can be assigned a value between 0 and 1. The role of β in this case is to control the significance of the average training image in the design of the filter. By controlling the value of β the filter is prevented from giving false detections due to the biased treatment of the low frequency components [74].

In order to clearly define the role of β the equation 3.22 the expression for *AICH* can be further simplified as given by equation 3.23 below [74].

$$AICH = h^{+}\left[\frac{1}{N}\sum_{i=1}^{N}(x_i - \beta m)\bullet(x_i - \beta m)^{+}\right]h \tag{3.23}$$

where equation 3.23 can be further simplified as given by equation 3.24 below [74]:

$$AICH = h^{+}C_x^{\beta}h \tag{3.24}$$

Where C_x^{β} is given by equation 3.25 below [74].

$$C_x^{\beta} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \beta m)\bullet(x_i - \beta m)^{+} \tag{3.25}$$

From the above *AICH* can be defined as the average measure of the correlation peak intensities of N images where each i^{th} training image has part of the mean subtracted which is given by $(x_i - \beta m)$. As discussed earlier, it is desirable for all the images in the training dataset to follow the behaviour of the average image. This can be achieved by

forcing every image in the training data set to have similar correlation output planes so as to match an ideal output plane given by u as given by equation 3.2 for MSE as discussed earlier. Now to find the image u that best matches the ideal output plane, its deviation from the ideal plane is minimized by substituting $g_i = (X_i - \beta M)h^*$ in equation 3.2 for MSE . It can be seen from h^* that the conjugated response of the filter has been considered. Now to find the optimal shape vector u_{ot} the gradient of MSE with respect to $u$is set to zero. This is further elaborated by equations 3.26 and 3.27 given below [74].

$$u_{ot} = \frac{1}{N} g_i \tag{3.26}$$

Now by substituting the value of g_i in equation 3.26 it becomes [74].

$$u_{ot} = (1-\beta)Mh^* \tag{3.27}$$

Now that the value of the optimal shape vector u_{ot} has been derived the ASM can be modified to calculate the dissimilarity of the training images to u_{ot} this new measure is known as Modified Average Similarity Measure ($MASM$) which is further elaborated by equation 3.28 given below [74].

$$MASM = h'\left(\frac{1}{N} \sum_{i=1}^{N} \left[X_i - (1-\beta)M \right]^* \left[X_i - (1-\beta)M \right] \right) h^* \tag{3.28}$$

where h' represents the transpose operation on the filter, equation 3.28 can be further simplified and re-written as given by equation 3.29 [74].

$$MASM = h^+ S_x^\beta h \tag{3.29}$$

where S_x^β is given by equation 3.30 [74].

$$S_x^{\beta} = \frac{1}{N}\sum_{i=1}^{N}\left[X_i - (1-\beta)M\right]^{*}\left[X_i - (1-\beta)M\right] \tag{3.30}$$

From the above it can be deduced that the *ASM* is a good measure to add distortion tolerance to the design of the filter but in the case where there are more than a certain number of training images having similar intensity level and low margin of change between the images then there is a reduction in the discrimination ability of the MACH to reject clutter. But on the other hand the newly defined measure *MASM* enables the filter to capture the finer details of the training images so that they can be classified easily and with fewer chances of errors. The inclusion of this measure makes the EMACH filter more sensitive to background clutter in the images.

In the case of the EMACH filter the *AICH* can be maximized and *MASM* minimized by controlling the parameter β. In an ideal case scenario it is expected to maintain a balance between the distortion tolerance and clutter rejection capabilities of the filter. So from the modified measures discussed above the transfer function for the EMACH filter can be given by equation 3.31 below [74]:

$$J^{\beta}(h) = \frac{AICH}{ONV + MASM} \tag{3.31}$$

By substituting the values of the measures in equation 3.31 [74]:

$$J^{\beta}(h) = \frac{h^{+}C_x^{\beta}h}{h^{+}\gamma Ih + h^{+}S_x^{\beta}h} \tag{3.32}$$

From equation 3.32 it can be seen that the *ONV* uses the additive white noise with variance γ. From the above equation it can also be seen that *ONV* helps to maintain noise tolerance when the value of β increases. By further simplification of equation 3.32 the following condition for the EMACH filter can be obtained [74].

$$(\gamma I + S_x^{\beta})^{-1}C_x^{\beta}h = \lambda h \tag{3.33}$$

In equation 3.33 the term λ is a scalar identical to $J^{\beta}(h)$ which shows that h is an eigenvector of $(\gamma I + S_x^{\beta})^{-1} C_x^{\beta}$ having the eigenvalue λ.Now it has been established that λ is the same as $J^{\beta}(h)$, so from this deduction it can be shown that $h$should be the eigenvector that corresponds to the maximum eigenvalue. The other eigenvectors provide smaller $J^{\beta}(h)$ non zero eigenvalues which may provide better performance due to the fact β is not known as a priori. The EMACH filter can be further expressed in the following terms [74].

$$h = \max - eigenvector\left\{(\gamma I + S_x^{\beta})^{-1} C_x^{\beta}\right\} \tag{3.34}$$

Now from equation 3.34 C_x^{β} and $(\gamma I + S_x^{\beta})^{-1} C_x^{\beta}$ are non-diagonal matrices for larger images a larger filter must be designed, but due to the non-linear nature of the filter a lot of computational resources and memory is required to avoid memory overflow problem.

There are various computational techniques that can be used to compute the transfer function for the EMACH filter but even these techniques require a lot of processing power and time. Although the EMACH filter provides better discrimination capabilities than the MACH filter it does take much more computational resources to design the filter.

For the scope of this thesis a filter is required that is robust and easy to design in real time i.e. the filter parameters can be changed in real time. But at the same time a filter is required that provides better discrimination ability to reject clutter and should be tunable to achieve a compromise for clutter rejection and discrimination abilities like the EMACH filter.

Hence for this reason an Optimal Trade-Off Maximum Average Correlation Height (OT-MACH) filter was considered for the scope of this thesis. OT-MACH filter is similar to the EMACH in many ways as it uses three parameters α , β and $\gamma$to control the measures of the MACH filter. The detailed design and capabilities of the OT-MACH filter implemented in the frequency domain are discussed in more detail in the forthcoming sections of this chapter.

3.4 Frequency Domain OT-MACH

In this section an OT-MACH filter has been designed in Frequency domain to experiment with the discrimination abilities to reject clutter.The major aim of the OT-MACH filter is to find the optimal compromise between good discrimination ability and distortion tolerance in the presence of noise. As discussed in the previous section the MACH filter maximises the relative height of the average correlation peak with respect to the expected distortions. However, the peak height of the MACH filter is unconstrained, making it more difficult to interpret the results of the correlation. As mentioned in Section 3.2, the MACH filter is derived by maximising a performance metric called the Average correlation height (ACH). However, several other performance measures have to be balanced to better suit different application scenarios. These measures are the Average Correlation Energy (ACE), Average Similarity Measure (ASM) and Output Noise Variance (ONV) [75] , [76].

The expression for the ASM has been given by equation 3.9 above and is the measure of distortions in the correlation plane relative to an average shape. Ideally all the correlation planes should be identical for distortion in-variant unconstrained correlation filters, and the value of ASM should be zero. Minimizing the value of ASM improves the stability of the filter's output response in the presence of distortions. As discussed above the MACH filter relaxes the correlation peak height constraints and maximizes the peak intensity of the average training image which has been achieved by using the measure known as ACH . This has been realised by maximizing the correlation height by correlating with the average of the training images and maximizing the magnitude of the resultant peak as given by equation 3.11 above. Also as discussed above in the design of correlation filters background noise plays a key role in contributing to the overall signal noise. Keeping that in mind variance ratios have been used in the design of the filter to make the filters more robust. One of these relations is called ONV which should be minimized as this is the measure of variance of noise in the correlation plane. The expression for ONV has been given by equation 3.12 above which is the filter's ability to reduce variance in noise at the output [9] , [68].

Another important criterion which has not been discussed earlier is the ACE .Which is the measure to assess the performance of the correlation filter and to provide a detailed idea of the energy present in the correlation output plane. In order for an optimal

correlation the value of ACE must be minimized to eradicate false detection of un-authorized classes. This can be further clarified by equation 3.35 given below [9], [77].

$$ACE = h^{+} D_x h \tag{3.35}$$

where D_x is the diagonal power spectral density of the training images and x_i is a diagonal matrix with elements of the training images along the diagonal as can be seen in equation 3.36 below.

$$D_x = \frac{1}{N} \sum_{i=1}^{N} X_i^{*} X_i \tag{3.36}$$

where in equation 3.36 X_i is diagonal matrix of the i^{th} training image. Now from the performance measures discussed above the concept of an optimal trade-off filter can be formed. To form an optimal trade-off filter, the following energy function is formed and minimised with respect to one criterion, holding the others constant as can be seen by equation 3.37 below [9].

$$J(h) = \alpha(ONV) + \beta(ACE) + \gamma(ASM) - \delta(ACH) \tag{3.37}$$

By substituting the values of the performance measures as discussed above in equation 3.37:

$$J(h) = \alpha h^{+} Dh + \beta h^{+} D_x h + \gamma h^{+} S_x h - \delta h^{+} m m^{+} h \tag{3.38}$$

From equation 3.38 the resulting transfer function for an OT MACH filter (in the frequency domain) has been formed [9]:

$$h = \frac{m_x^{*}}{\alpha C + \beta D_x + \gamma S_x} \tag{3.39}$$

From equation 3.39 an OT-MACH filter is given where α, β and γ are the non-negative parameters, m_x is the average of the training image vector $x_1, x_2, ..., x_N$ (in the frequency

domain), and *C* is the diagonal power spectral density matrix of additive input noise. It is usually set as the white noise covariance matrix, $C = \sigma^2 I$. Also D_x is the diagonal average power spectral density of the training images as given by equation 3.36 above.

S_x denotes the similarity matrix of the training images which has been shown by equation 3.10 given above [9].

In the case of the OT-MACH filter the different values of α, β and γ control the filter's behaviour to match different application requirements. If $\beta=\gamma=0$, the resulting filter behaves much like a MVSDF filter with relatively good noise tolerance but broad peaks. If $\alpha=\gamma=0$ then the filter behaves more like a MACE filter which generally exhibits sharp peaks and good clutter suppression but is very sensitive to distortion of the target object. If $\alpha=\beta=0$, the filter gives high tolerance for distortion but is less discriminating [9], [40].

Due to its configurable capabilities and relatively faster filter design process in terms of computations as compared to the EMACH, the OT-MACH was chosen for the scope of this thesis so as to experiment in situations where filter should be designed or configured in real time. In order to assess the abilities of the OT-MACH filter it should be tested in different scenarios and for part of that purpose a rigorous test case has been created. An OT-MACH filter designed in frequency domain has been tested in different situations which are discussed in more detail in the forthcoming section of this chapter.

3.4 Performance Analysis of the OT-MACH filter

As discussed earlier on that the capabilities of the OT-MACH filter can be tuned to achieve a balance in noise tolerance and discrimination abilities. In this section the OT-MACH has been tested in terms of experimenting with images containing background noise and images without background noise. Different orientations have been introduced from the training images and their effects have been noted on the overall correlation plane by mean of performance ratios like the COPI, PCE and PSR discussed in the previous chapter. The main purpose has been to assess what advantages an unconstrained correlation filter like the OT-MACH gives over the constrained correlation filter SDF as discussed in the Chapter 2.

3.4.1 OT-MACH Experiments without Background Noise

In order to use the OT-MACH filter the first step is to set the performance parameters to an optimal value. There are many ways of defining the values of the parameters for the OT-MACH but the most commonly defined rule from trial and error is to keep the value of β high between a range of 0-1 and the value of γ as low as possible.

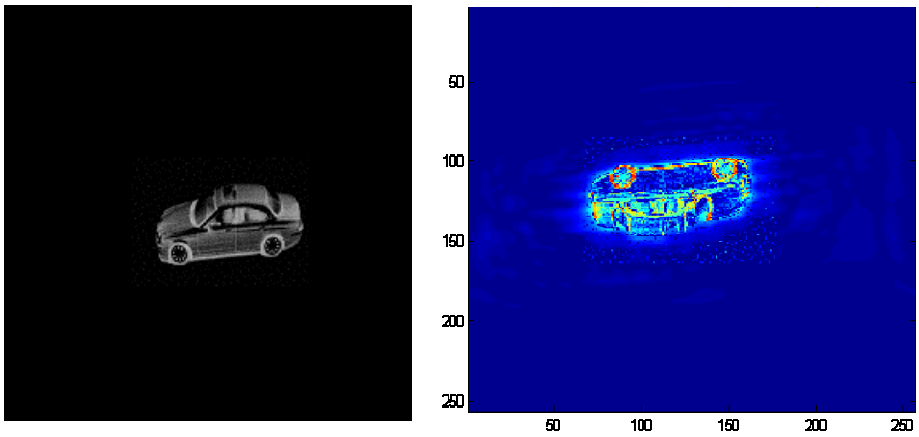
The performance parameters used in the OT-MACH filter function were set to α = 0.01, β = 0.6 and γ = 0.1. The α value was set as low as possible to produce sharp correlation peaks. The β value was made large to produce sharp peaks but without resulting in poor distortion tolerance, which is directly related to the difference in consecutive training set orientation angles. The γ value was set low to maximise the filter's discrimination ability between in-class and out-of-class objects

The first test carried out was a basic correlation using one of the training set images which is given by figure 2.3 in Chapter 2 as the input image. It was expected that the filter would perform well in this test since the input image was directly associated with the construction of the filter. The test was therefore similar to autocorrelation.

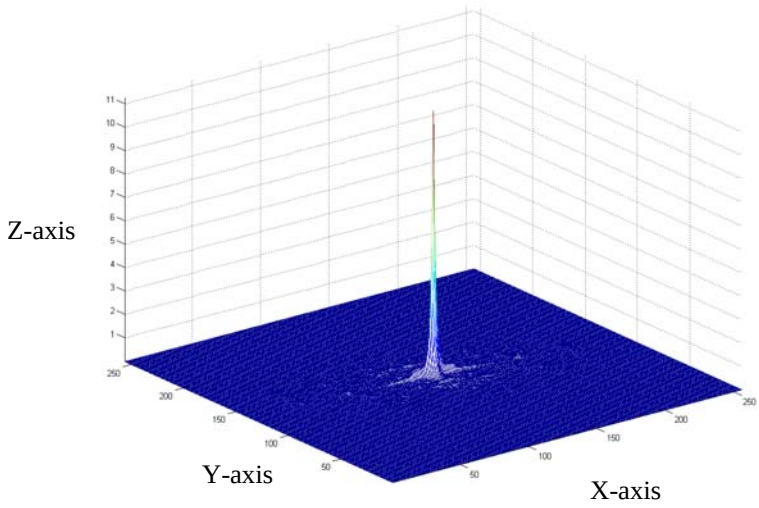
The output results for the OT-MACH filter produced a sharp peak and no out-of-class peaks. The average peak-to-correlation energy (PCE) value of the output correlation planes was 0.01, which shows a high detection capability with a low chance of false detections caused by false peaks. The correlation plane shown is where the strongest peak is found during scanning of the image using the moving window method described above i.e. centered over the target object.

The initial experiments were conducted in the presence of no background noise in order to evaluate the performance of the filter. The main purpose was to enable the filter to perform independently and to determine a set of parameters which would enable invariance to change in orientations of the target and yet discriminate in the presence of background noise.

In Figure 3.1 it can be seen that when a reference image of a car image oriented at 0 degrees is correlated with another training image of a car at the same angle of rotation without any background noise.



(a) **(b)**



(c)

Figure 3.1 Car reference image (a) , Filter Impulse response (b) and Output Plane (c)

This can also be used to show the filter's impulse response as both of the images are the same. In this ideal case the values for the performance ratios achieved can be seen from Table 3.1 below:

PCE	PSR	Orientations (Target , Reference)
0.057	146.59	(0,0) + No Background Clutter

Table 3.1 PCE\PSR at zero degree orientation

From Figure 3.1 it can be seen that the OT-MACH filter produces a much sharper peak than the SDF filter discussed in the previous chapter.

Next a test of the filter correlating a car image at 0 degree orientation with a 10 degree training image orientation was made , the result of which can be seen from Figure 3.2 below:

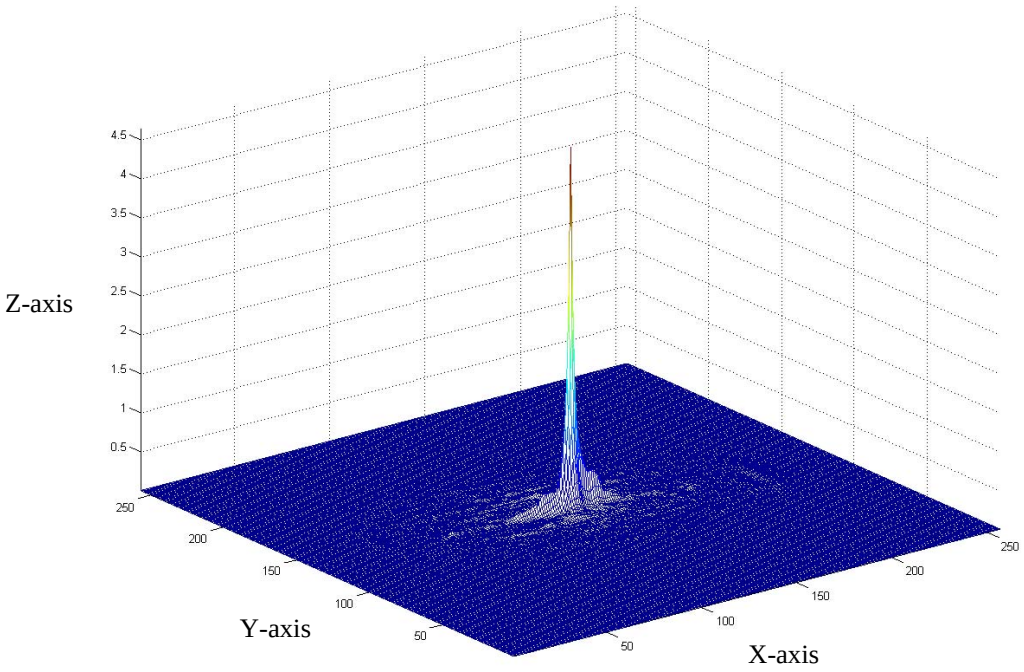


Figure 3.2 Correlation with the filter trained at 10 degree with no background clutter

The results of the correlation can be seen from Table 3.2 given below:

PCE	PSR	Orientations (Target , Reference)
0.03	120.59	(0,10) + No Background Clutter

Table 3.2 PCE\PSR at 10 degree orientation

It can be seen that there is slight drop in the PCE and PSR values this is due to the fact that the filter is being tested on a different orientation and hence being correlated with the same target but a different degree of orientation. Even in this scenario the filter tends to give a good correlation peak exhibiting the filters ability to match against different orientations.

Now in order to assess the ability of the OT-MACH filter in frequency domain a series of tests have been conducted between different set of orientations ranging from 0-70 degrees.

From figure 3.3 it can be seen that when the target image oriented at 10 degrees is correlated with a filter trained on three orientations ranging from 0-30 it was observed that even with the increased orientations in the design of the filter a match was made.

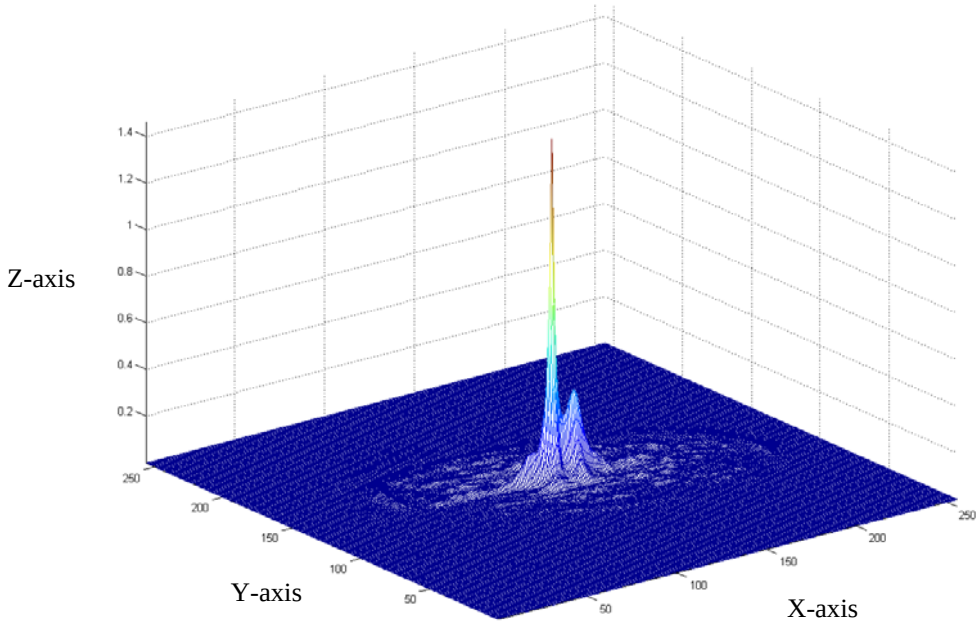


Figure 3.3 Correlation with the filter trained at 0-30 degree with no background clutter

PCE	PSR	Orientations (Target , Reference)
0.0141	93.22	(0,0-30) + No Background Clutter

Table 3.3 PCE\PSR at 0-30 degree orientation

From Table 3.3 the performance ratio values for Figure 3.3 can be seen. It can also be seen from Figure 3.3 that a visible side lobe has been formed as well which contributes to the overall reduction of PSR value and could lead to false detection of targets. Also there is a drop in the height of the peak, as indicated by the reduced PCE value.

From Table 3.3 it can be seen that adding extra orientations into the design of the filters improves the detection ability but reduced the overall energy of the correlation plane. Now in order to understand the extent of the detection capabilities of the filter it has been trained with orientations from 0-70 degrees and correlated with a target image at 0 degree orientation. This can be seen from Figure 3.5 below:

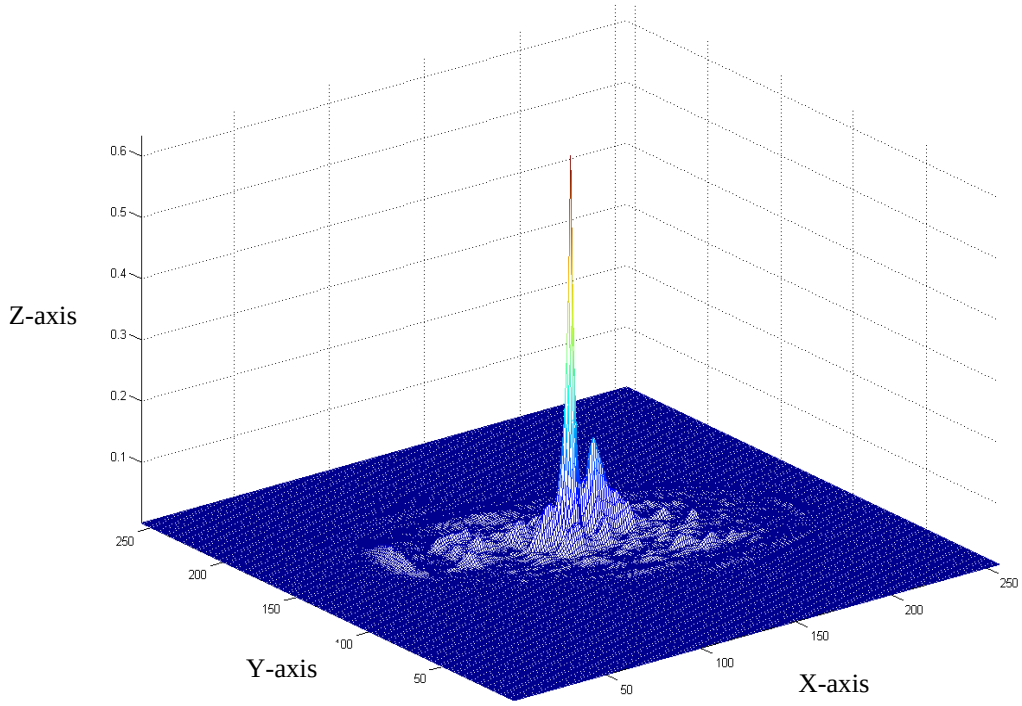


Figure 3.4 Correlation with the filter trained at 0-70 degrees with no background clutter

It can be seen from Figure 3.4 that there is a further decrease in the height and sharpness of the peak as well as the formation of a larger side lobe increasing the chances of a false detection. There is an ideal compromise here as by increasing the number of training image the filters ability to detect the target are increased but its ability to discriminate against background noise are reduced at the same time.

A series of experiments were conducted for the work of this thesis to assess the capabilities of the OT-MACH filter in frequency domain. Initially the PCE values where considered to assess whether a mean threshold could be calculated or not. The results of which can be seen from Table 3.4 given below:

PCE Value	Car Orientation
0.057	0
0.03	0,10
0.0196	0,10,20
0.0141	0,10,20,30
0.0106	0,10,20,30,40
0.0094	0,10,20,30,40,50
0.0083	0,10,20,30,40,50,60
0.0074	0,10,20,30,40,50,60,70
Threshold	0.01955

Table 3.4 PCE values for OT-MACH trained at multiple orientations

From Table 3.4 it can be deduced that since OT-MACH is an unconstrained filter it is very difficult to assess the quality of the peak. So in order to define the quality of detection a threshold has been calculated. The threshold value given in Table 3.4 has been calculated by taking the mean of all the PCE values at multiple orientations.

This shows that the performance of the filter is degraded after more than three orientations were included in the design of the filter. It can be deduced that when the correlation peak height was below the mean threshold the quality of detection was not satisfactory. In order to maximize the filter performance we must thus re-tune the filter parameters. But the only problem with this approach is that the variation in the PCE is not linear so a justified performance analysis cannot be deduced. A more viable approach was to include the PSR metric in the table to assess the performance of the filter.

By assessing the results from the previous experiments and keeping into account that there were side lobes being formed the new parameters are set accordingly to reduce the side lobes. The new parameters have been changed from α = 0.01, β = 0.9, γ= 0.0001

to α = 0.001, β = 0.6 and γ= 0.001. With the new parameters there was a significant improvement in the detection abilities of the filter.

A new set of experiments were conducted on the same pattern as conducted earlier on this section without the presence of background noise. The results for the experiments with the new parameters can be seen from Table 3.5 given below:

PCE	PSR	Orientations
0.1685	202.06	0
0.0946	187.26	0,10
0.686	184.15	0,10,20
0.0498	174.39	0,10,20,30
0.0357	164.45	0,10,20,30,40
0.0318	161.29	0,10,20,30,40,50
0.0291	158.86	0,10,20,30,40,50,60
0.0269	156.67	0,10,20,30,40,50,60,70

Table 3.5 PCE\PSR values with new parameters

It can be seen from Table 3.5 that the new set of PCE and PSR values are considerably higher than the previous values from Table 3.4. In this case the correlation threshold is set at 0.02 which is based on selecting the worst case scenario in reference to PCE. Based on the new set of parameters the filter was trained and a further set of experiments have been conducted with the inclusion of background noise to assess the capabilities of the filter.

3.4.2 OT-MACH experiments with background noise

As discussed earlier the detection capability of the filter must be tested with background noise present as this will define the discrimination abilities of the filter. Also the degree of invariance to distortion and orientation will be established. In this case a compromise between the distortion tolerance and clutter resistance capabilities of the filter will then be formed.

A set of results have been generated with background noise present containing the target oriented at 0 degrees. The filter has been trained using the same parameter values as discussed in section 3.4.1. The target image with background noise and the output plane can be seen from Figure 3.5 given below:

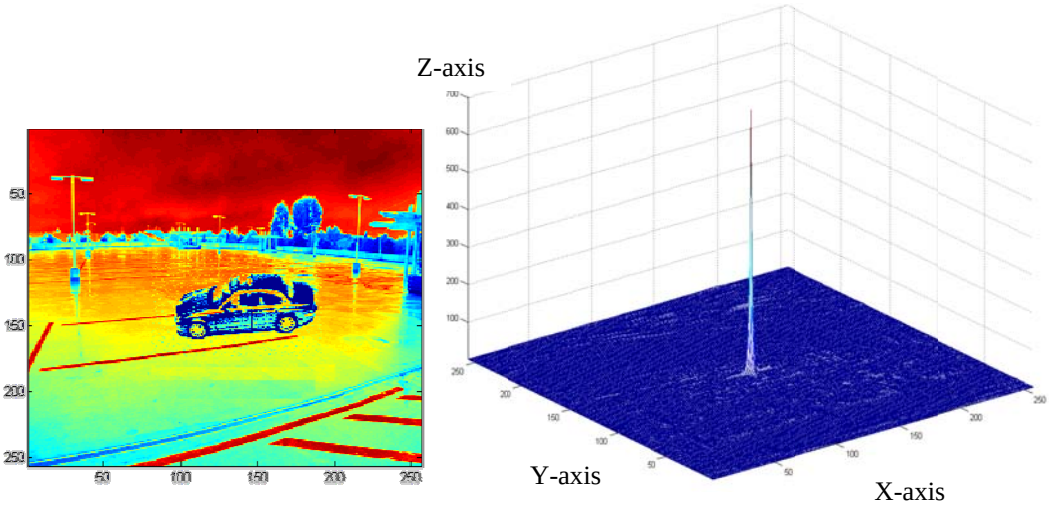


Figure 3.5 Target image with background noise (a), output plane for OT-MACH filter trained at 0 degree orientation (b)

From Figure 3.5 it is shown that when a target image containing a car oriented at zero degrees with background clutter. The reference image is a car at zero degrees with a black background upon which the filter is trained. As can be seen from Figure 3.5b a very sharp peak is obtained, in this case the PCE of the correlation plane is 0.0276 and the PSR is 193.99. This shows that the filter is successfully discriminating against background noise and giving a good detection peak.

However, this is the case when only one orientation is used. When the filter has been trained using orientations from 0-70 degrees the correlation output can be seen from Figure 3.6 below.

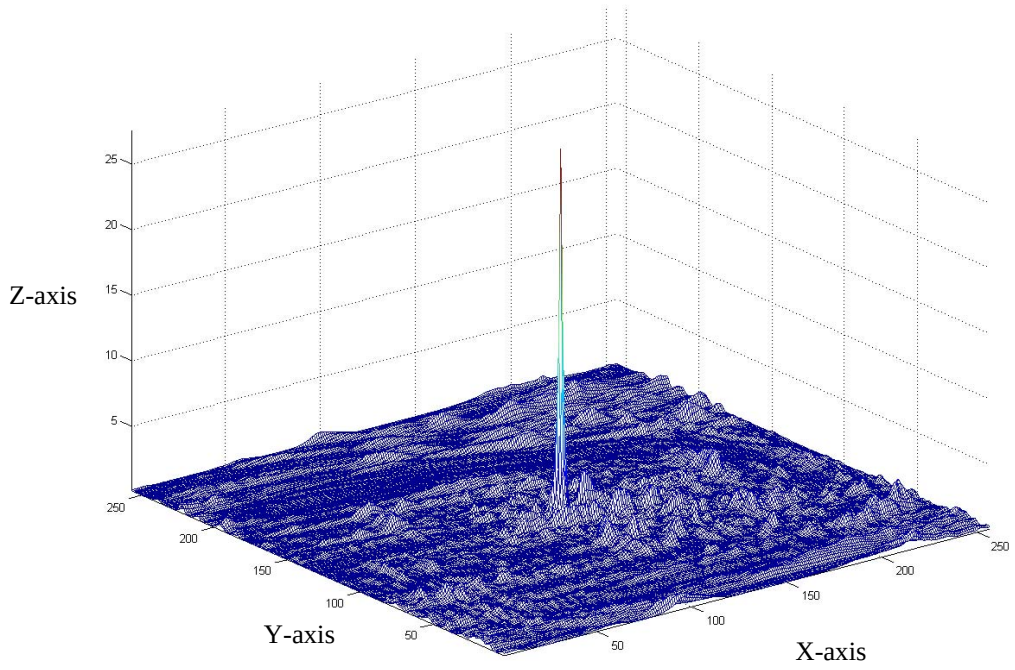


Figure 3.6 Correlation with the filter trained at 0-70 degrees with background noise

In this case from Figure 3.6 it can be seen that a sharp peak is present but there is a lot of background noise in the correlation plane. The PCE = 0.0276 and the PSR = 105.52 and, although the peak is sharp, the inclusion of a lot of background noise and side lobes increases the chances of false detection. So the parameters for the OT-MACH filter were changed again to improve the discrimination capabilities.

The value of the parameters were changed to α = 0.0001, β = 0.7, γ = 0.001. The value of β has been increased to control the discrimination ability and α has been reduced to suppress the background noise. The new correlation output plane can be seen from Figure 3.7.

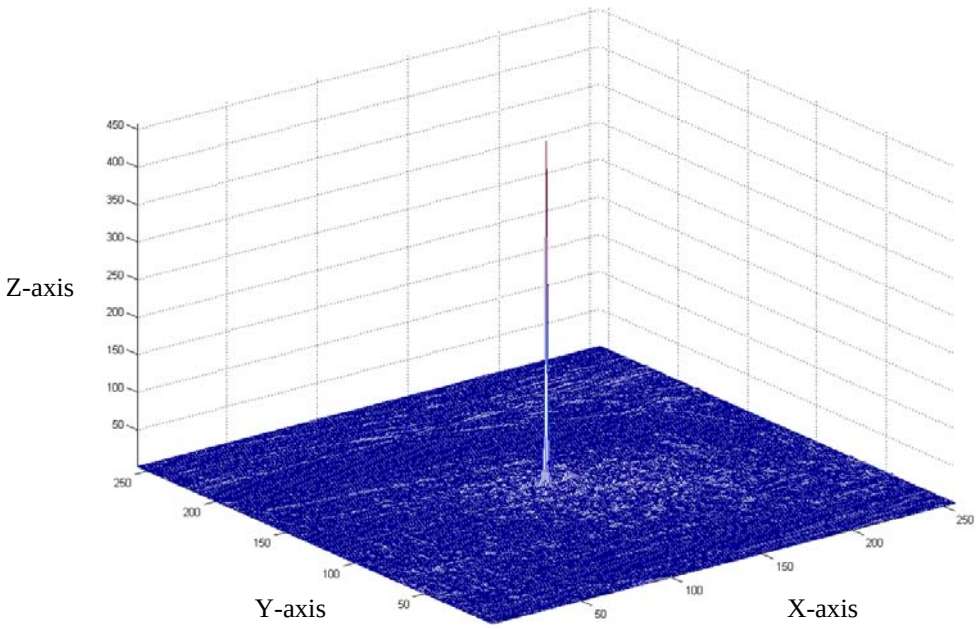


Figure 3.7 Correlation with the filter trained at 0-70 with background noise

In Figure 3.7 the new correlation output plane achieved from the new set of parameters has a sharper peak and less background noise having the PCE = 0.0192 and PSR = 231.38. The re-tuning of filter has made the filter more invariant to the background noise.

In a practical implementation, the filter would be expected to correctly detect in-class targets even if they were at an angle that did not exactly match any of the training set images. This is because the filter can only be constructed with a training set consisting of a discrete set of distortion variations, whereas the input images can have any continuous distortion variation. To test the distortion tolerance of the filter it was correlated with a set of images of the car whose angle fell between that of the training set images.

The filter was trained on eight reference car images with varying orientation from 0 to 70 degrees in steps of 10 degrees with black background. It was then correlated with an input image which had a 5 degree orientation difference from the most similar training set images i.e. at an angle intermediate to two training images. The result can be seen from Figure 3.8 given below.

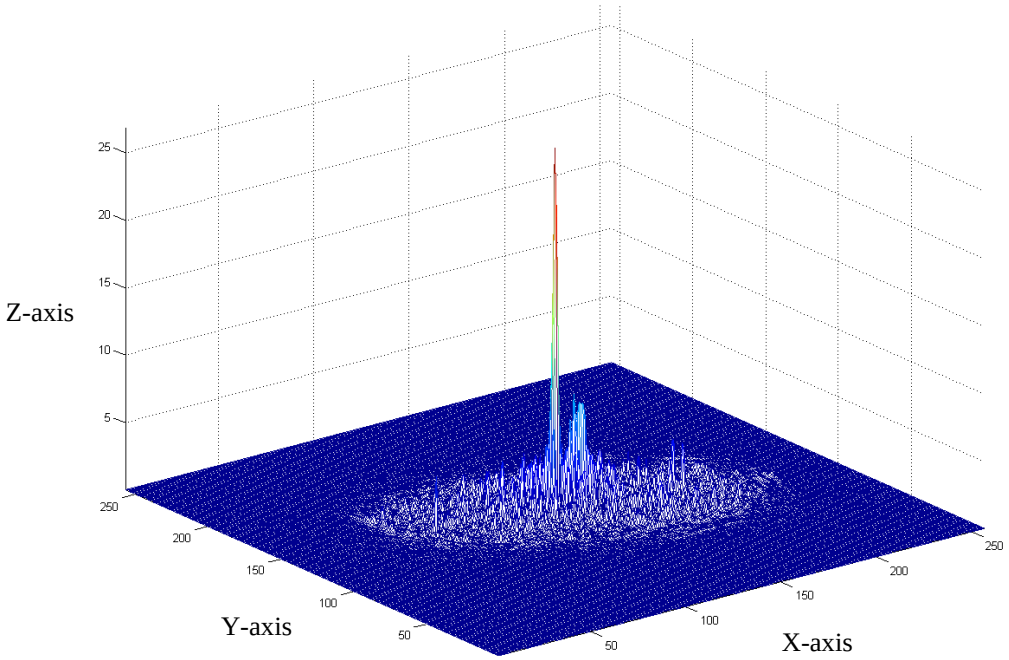


Figure 3.8 Correlation with the filter at 0-70 degrees with target at intermediate orientation

As expected, the result shows a small average reduction in COPI and a slight broadening of the base of the peak; however, the OT-MACH filter's response was still sharp and well above the detection threshold giving an average PCE value of 0.009 and PSR of 84.3. This demonstrates that the filter can still easily maintain its discriminating ability despite being correlated with input images at intermediate angles to the reference images used to construct the filter. This ability means that the filter can cope well if only a coarse set of reference images is available with only a small number of images covering a large range of distortion [62].

Now from the above set of conducted experiments it was observed that the filter performed at its best when correlated with input images in the middle of the range of reference images used to build the filter. For example, the above filter gave higher COPI values when correlated with reference images at orientation angles of 10 and 20 degrees rather than when correlated with images at angles of 0 and 30 degrees. This can be explained due to the way the MACH filter combines the reference images by calculating their mean. The resulting filter is slightly biased towards the middle of the range of orientations since the images at the edges of the range tend to be less well presented. This effect would be removed if the range of distortion wrapped round; for example, if a range of orientation angles was used that equally sampled the maximum range of 360 degrees.

The OT-MACH filter designed in the frequency domain has been shown to have discrimination abilities to background noise in the above experiments. However there is a limitation to the capabilities of the OT-MACH which is further elaborated in the next section.

3.5 Limitations of OT-MACH in Frequency Domain

As discussed in Chapter 1 the defined problem area for OT-MACH is the filter's inability to discriminate in non-uniform brightness hence making it incapable illumination invariant. In order to illustrate the inability of the frequency domain OT-MACH to be illumination invariant, the test shown by Figure 3.9 was conducted.

In this test a brightly illuminated square having pixel intensity of 250 is used to train the frequency domain OT-MACH filter and the target image contains squares at different illuminations ranging from 40 to 255 which is shown by Figure 3.9.

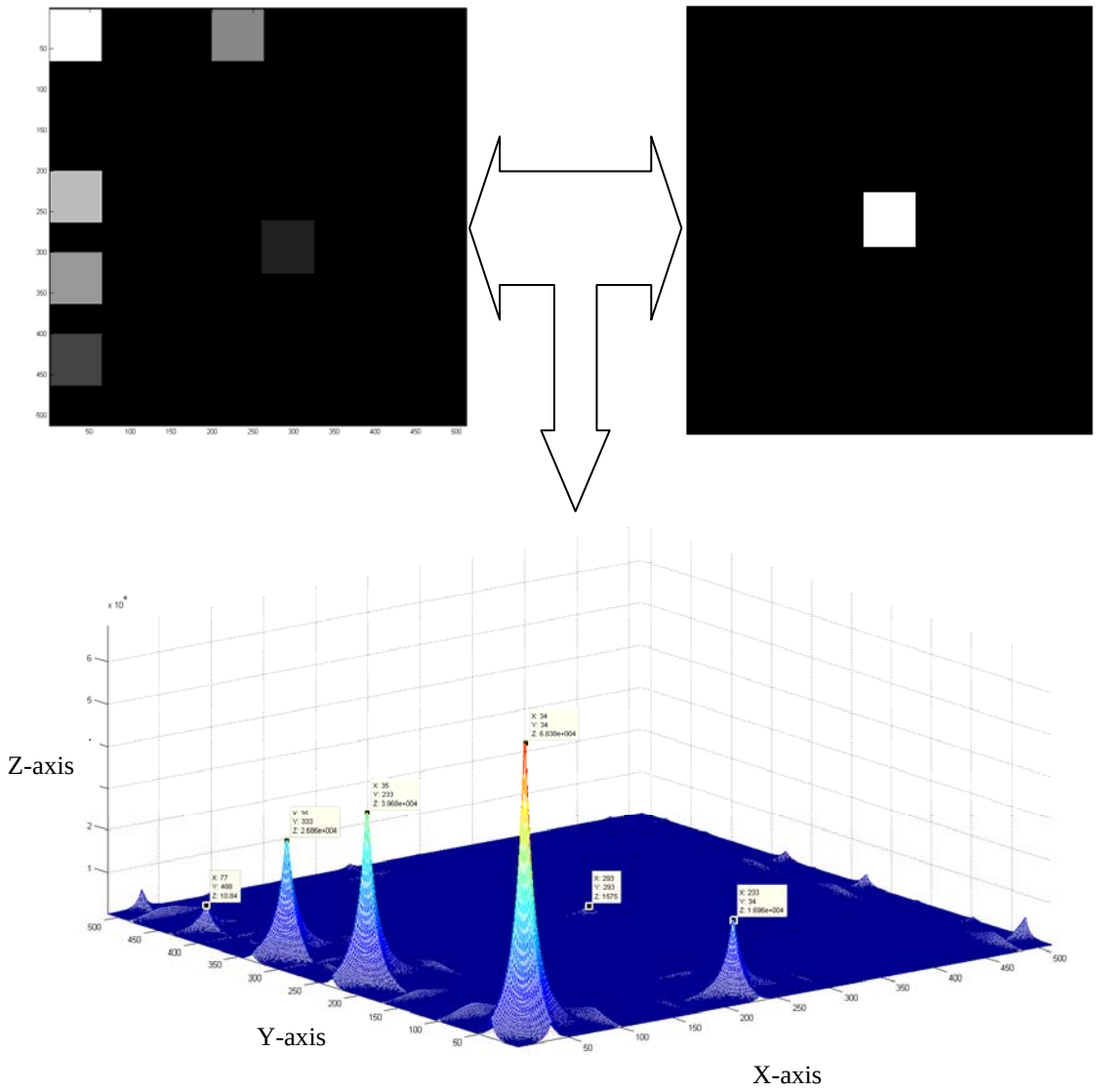


Figure 3.9 Limitation of OT-MACH in Frequency Domain

From Figure 3.9 it can be seen that there are multiple peaks in the output plane. It can be seen that the square with the brightest illumination has the highest peak in the corner and as the pixel intensity decreases the peak size reduces. And if observed closely the frequency domain OT-MACH fails to pick up the square having the pixel intensity of 40 located in the center of the target image. This shows the frequency domain OT-MACH's inability to detect object if it is in a dark region although being the same objet.

This is due to the fact that the energy is not localized to certain area but rather distributed equally on the whole correlation plane. In order to overcome this problem the energy should be localized to a window, for this reason a Spatial Domain approach

has been proposed to implement the OT-MACH instead of frequency domain in this thesis which uses a local window energy normalisation to overcome this problem. This has been discussed in more detail in Chapter 4.

3.6 Conclusion

In this chapter the MACH filter and its enhancements have been summarised. It has been shown that MACH filter can be set to find the optimal compromise between tolerance to target distortion and discrimination in the presence of background noise. To achieve this, the MACH filter maximizes the relative height of the average correlation peak with respect to the expected distortions included in the design of the filter. Unlike the SDF, the MACH filter can be tuned using the weighted coefficients in its transfer function to maximize the correlation peak height, peak sharpness and noise suppression whilst also being tolerant to distortions in the target image.

In this chapter, a brief discussion of the origin of correlation filters and different correlation techniques has been presented. The frequency domain implementation of the OT-MACH has been discussed in detail and the discrimination abilities with and without background noise has been tested. A set of results has been produced which are quantified using PCE and PSR measures. A tabular set of results has been produced to show that there is degradation in performance of the OT-MACH filter with the addition of multiple target orientations and how the filter response can be improved by tuning the OT-MACH filter parameters. Also, the limitations of the OT-MACH filter, when implemented in the frequency domain, in its inability to detect targets in non-uniform brightness has been shown and a spatial domain implementation to overcome this problem has been proposed. Additional capabilities of the spatial domain implementation of the filter will be discussed in Chapter 4.

Chapter 4

Spatial domain Optimal Trade off Maximum Average Correlation Height (SPOT-MACH) filter

4.1 Introduction

A frequency domain implementation for the OT-MACH filter was discussed in Chapter 3 and a brief overview of its limitations was also given.

In Chapter 3 it was shown that when employing a frequency domain Optimal Trade-off Maximum Average Correlation Height (OT-MACH) filter degradation in performance could be seen when correlations were performed between a uniformly illuminated training dataset and a series of targets having variable illuminations.

In Figure 3.9 in Chapter 3 it was shown that when the pixel intensity of the training set was taken to be a of value 250 and correlated with a target scene having the same object as the reference image but at different illuminations, a varied set of peaks was obtained. In this case the object having the brightest pixel intensity gave the largest peak and the quality of the peak degraded with the reduction in the illumination. In an ideal case scenario the OT-MACH filter should have given an equal set of peaks for all targets regardless of their illumination. But not only does the OT-MACH filter, when implemented in the frequency domain, fail to distinguish the desired class but also gives two false peaks having a higher peak height than the peak heights achieved from the autocorrelation of the training images. This shows the extent of variation in the quality of the detection some changes in illumination could cause to the performance of the correlation filter.

As discussed in Chapter 2, the average training image used in the design of the MACH filter is beneficial in representing the average behavior of the desired class, but fails to capture the finer details of the target class. In most cases the average training images tend to look like cluttered images and this may lead to the MACH filter's inability to discriminate the desired class from cluttered backgrounds and hence increase false detection rates.

In order to further illustrate the degradation in performance of the frequency domain OT-MACH filter in non-uniform illumination, a set of experiments have been conducted using a brightly illuminated false target.

In this case the 512 x512 target scene consisted of two targets: a large triangle having pixel intensity of 250; and a 64x64 square having pixel intensity of 40. The target scene can be seen from Figure 4.1 given below.

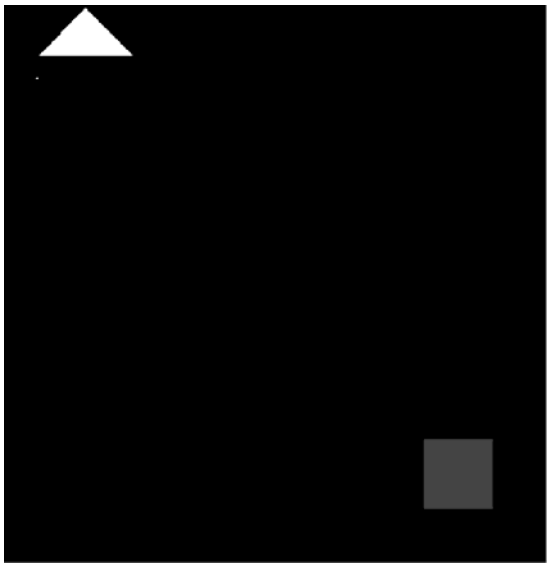


Figure 4.1 Target image for frequency domain OT-MACH illumination in-variance test

The OT-MACH filter in frequency domain was trained using a 512x512 reference image consisting of a square of 64x64 having a pixel intensity of 250 in the center of the image. The filter is kept simple with the inclusion of only a single image in the training set rather than multiple images in order to check illumination invariance rather than orientation. This can be seen in Figure 4.2 given below.

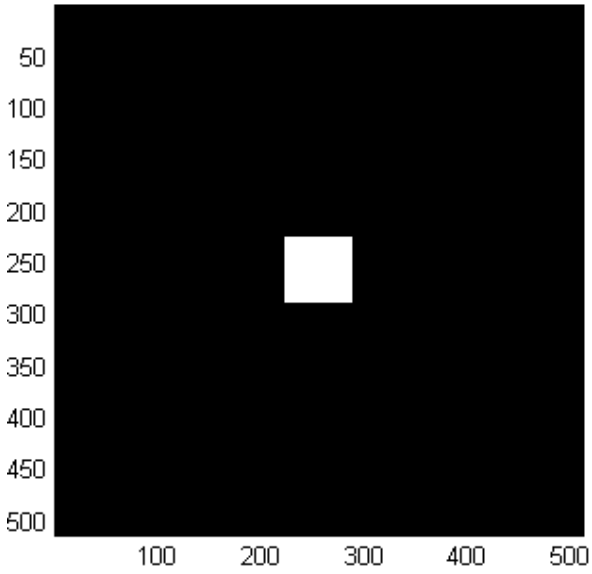


Figure 4.2 Training image for frequency domain OT-MACH

When the target image from Figure 4.1 and training image from Figure 4.2 were correlated the filter output plane can be seen from Figure 4.3 given below.

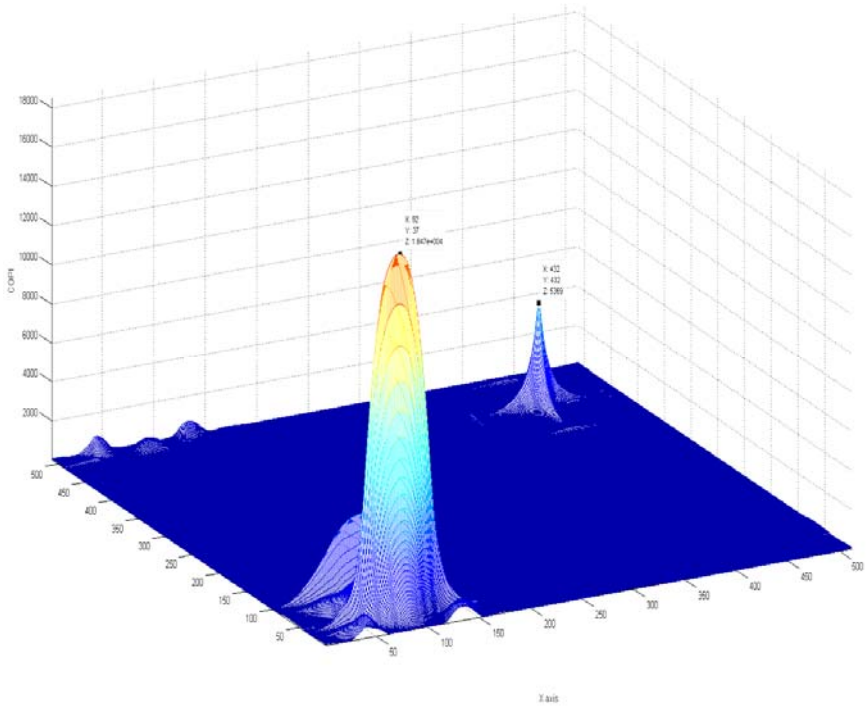


Figure 4.3 Output plane for OT-MACH in frequency domain

From figure 4.3 it can be seen that the *COPI* value for the false target which is the triangle is 1.847×10^4, and *COPI* for the true target which is the less illuminated square is only 0.5369×10^4. The OT-MACH implemented the in frequency domain was tuned with the following parameters $\alpha=0.8, \beta=0.9$ and $\gamma=0.7$ which were chosen by exhaustive hit and trial to find the most optimal values of the correlation plane.

This shows that the true target has a smaller correlation peak than the false target in the correlation output plane. This signifies the fact that the OT-MACH has given a false detection of target in the presence of non-uniform illumination

In order to enhance the detection capabilities of the OT-MACH filter in the frequency domain and to give an accurate detection of the target, the filter has been tuned to increase its discrimination capabilities so that the true target could be detected effectively in the presence of varying illumination. In this case the OT-MACH filter is tuned with the following parameters $\alpha=0.001, \beta=0.7$ and $\gamma=0.01$. The performance of the frequency domain OT-MACH filter can be seen from Figure 4.4 below.

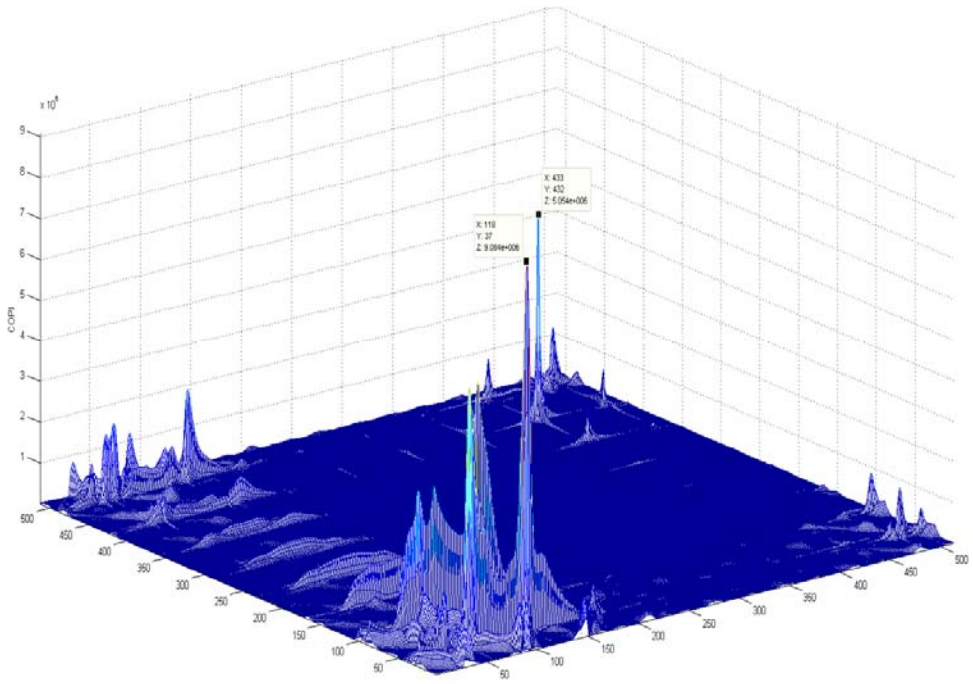


Figure 4.4 Frequency domain OT-MACH filter with changed filter parameters

From Figure 4.4 it can be seen that the *COPI* for the false target is now 9.084×10^{6} and for the true target has been given by 5.054×10^{6}. It can be deduced from this that even with the improved discrimination capabilities the OT-MACH when implemented in frequency domain was unable to distinguish between true and false targets. Although higher *COPI* values were achieved in this experiment but still the false target peak was higher than that of the true target.

In order to further investigate the capabilities of the OT-MACH when implemented in the frequency domain and to improve the chances of detection, it has been decided to test the detection capabilities of OT-MACH using training images having the same illumination as the target. In this case the filter was trained using a 512x512 image having a 64x64 square at the center with pixel intensity of 40 which is same as one of the targets. Also changing the parameters of the filter to the initial parameters values which were $\alpha = 0.8$, $\beta = 0.9$ and $\gamma = 0.7$. The new training image used to train the OT-MACH filter in frequency domain can be seen from Figure 4.5 below.

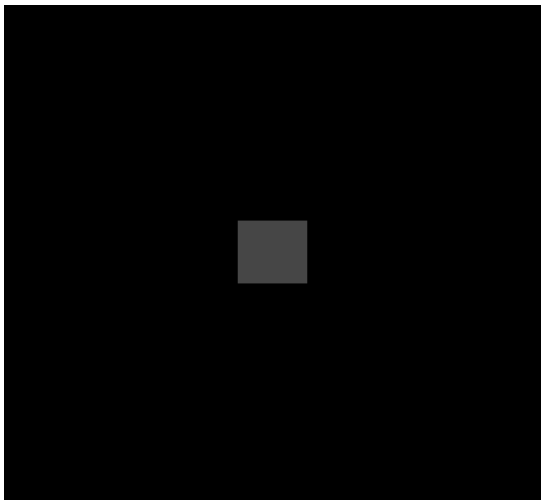


Figure 4.5 Modified reference image for OT-MACH when implemented in the frequency domain

The correlation output plane for the filter using the training image show in Figure 4.5 and the target image shown in Figure 4.1 is shown below in Figure 4.6.

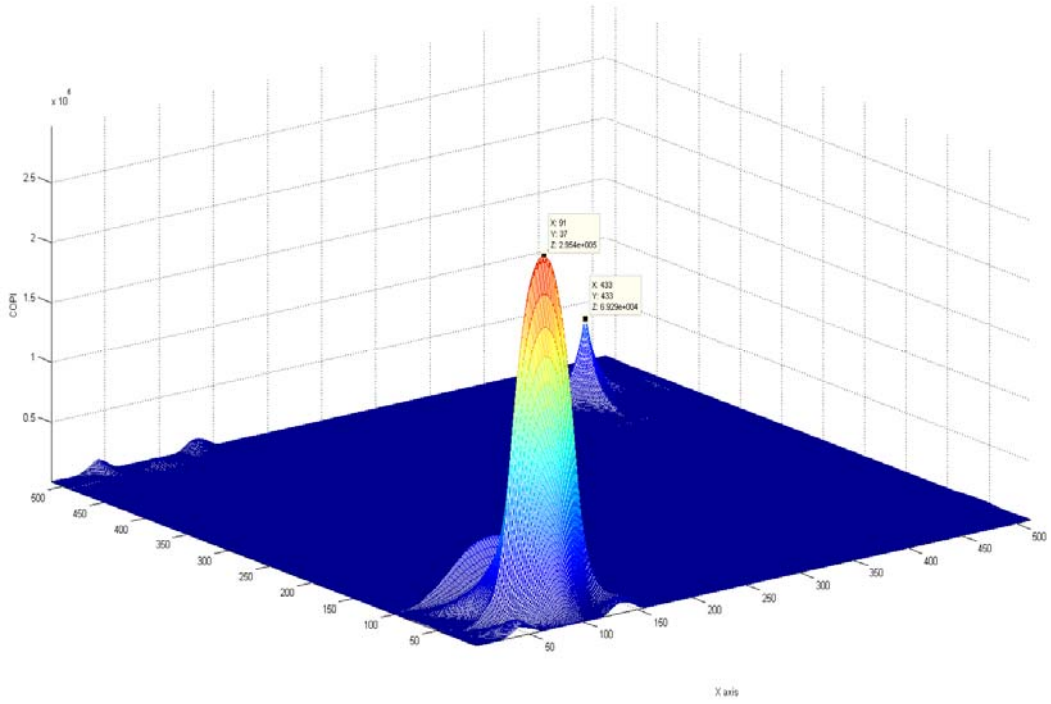


Figure 4.6 Output plane for frequency domain OT-MACH trained with new dataset

From Figure 4.6 it has seen that the *COPI* for the false target is 2.954×10^5 and for the true target is 6.920×10^4 which shows that even in the scenario where the training image dataset was at the same pixel intensity as of the target image the frequency domain OT-MACH filter failed to detect it.

This signifies the problem associated with the frequency domain implementation of the OT-MACH filter which is due to the fact that the energy of the reference and target objects is distributed globally over the frequency plane. In the case where there is a brightly illuminated object present in the scene, the correlation output values will be higher for that region and hence will produce a taller peak which results in a false detection.

In order to overcome this problem a spatial domain approach has been proposed for the implementation of the OT-MACH filter which employs local window energy normalization for the filter to facilitate more tolerance to illumination changes. The

proposed filter is named the Spatial Domain OT-MACH (SPOT-MACH) and its detailed design and analysis is discussed in more detail in the forthcoming sections of this chapter.

4.2 Why Spatial Domain Correlation Filters?

As has been discussed earlier in Chapter 2, it is necessary for a correlation filter to be invariant to all distortions of the target object while still being able to maintain a good discrimination between similar objects. In other words, the correlation filter must strike a compromise between the requirements of in-class distortion tolerance and out-of-class discrimination. A band-pass filter is ideal in this case where the high frequency response is attenuated to provide less sensitivity to in-class variations while low frequencies are also removed as they compromise the discrimination abilities of the filter.

The OT-MACH filter can be tuned accordingly to act as a band-pass filter to attenuate the desired range of frequencies. It has been shown in section 4.1 for a frequency domain implemented OT-MACH filter that global illumination normalization will not work on a scene containing a brightly illuminated false target. If the target is present in the shaded areas of the scene, and structured clutter is present in the illuminated portions then we may get a higher correlation peak for the structured clutter as compared to that of the real target. This makes the OT-MACH filter variant to changes in illumination and so does not satisfy the criterion of an ideal correlation filter it should be made tolerant to changes in the illumination. In this thesis it has been proposed that spatial domain implementation should be the preferred domain when designing a filter and local window energy normalization should be employed to provide illumination change tolerance.

In the case of the human visual system, the brightness is extracted from the surfaces of the visual scenes. Determining the brightness of the scene plays an important role as it influences the performance of the vision system when detecting shape or an object from a shaded area. In the case of object recognition an object must be recognized independently of its illumination. Also it is known that the perceived brightness of a surface depends on the brightness of the neighboring surfaces. In the case of

determining the extent of illumination present in a scene, this is not limited to measuring the light reflected from each surface since the surrounding surfaces influence the perceived brightness as well.

When employing the frequency domain implementations, researchers have tried to tackle the problem of illumination changes by using Oriented Difference-of-Gaussian (ODOG) filters and then applying global response normalization to equalize the amount of energy at each orientation across the entire scene [78]. The ODOG model has been extended by exploring the normalization schemes that can be employed within the image datasets. It has been determined that when using images having a non-uniform distribution of energy at different orientations, normalisation does not improve the performance of the filter. Furthermore, the filter fails completely in the scenario where there is equal energy at most of the orientations when integrated across the entire image. For this situation, it has been shown that global normalization techniques would even fail in cases where there is relatively uniform distribution across each

The failure of ODOG shows that in the frequency domain, normalization of energies across the scene is not beneficial. Also, the manipulation of spatial frequencies in order to normalize them is a complex process. In contrast, pixel intensities are straight forward to manipulate and provide a point to point indication of the change of energy across the scene. In this section, a spatial domain approach, which we call the SPOT-MACH, is proposed in this thesis which uses local window energy normalization to overcome the illumination variation problem.

4.3 Design of the SPOT-MACH filter

A moving window is used to implement a SPOT-MACH filter which can be locally modified depending upon its position in the input frame. This enables adaptation of the filter dependant on locally variant background clutter conditions and also enables the normalisation of the filter energy levels at each step. Thus the spatial domain SPOT-MACH filter offers an advantage over its frequency domain implementation as shift invariance is not imposed upon it. The only drawback of the spatial domain implementation is the amount of computational resources that are required for its real time implementation. Recently an optical correlator using a scanning holographic

memory has been proposed by Birch *et al* for the real time implementation of space variant filters of this type [79].

A detailed consideration of the proposed SPOT-MACH filter is presented in the forthcoming sections of this chapter. A general overview of the working of the SPOT-MACH can be seen from Figure 4.7 which shows that the SPOT-MACH has two inputs, one the Input Image, which the target image is having the background clutter in most scenarios, and the Reference Image. Both of the images are passed through the SPOT-MACH transfer function to create a composite image and provide edge enhancement to the target image. The spatial correlator extracts the window from the kernel image depending upon its size and then applies point to point normalized cross-correlation with the target image. The spatial correlator calculates the maximum peak which can be used to give the position of the object. The decision matrix then checks if the maximum peak is equal to, or above, a threshold which can be set to be up to eighty percent of the value of the maximum auto-correlation intensity. The numbers of iterations of the window are counted and the algorithm does not terminate until the window has scanned the full target image.

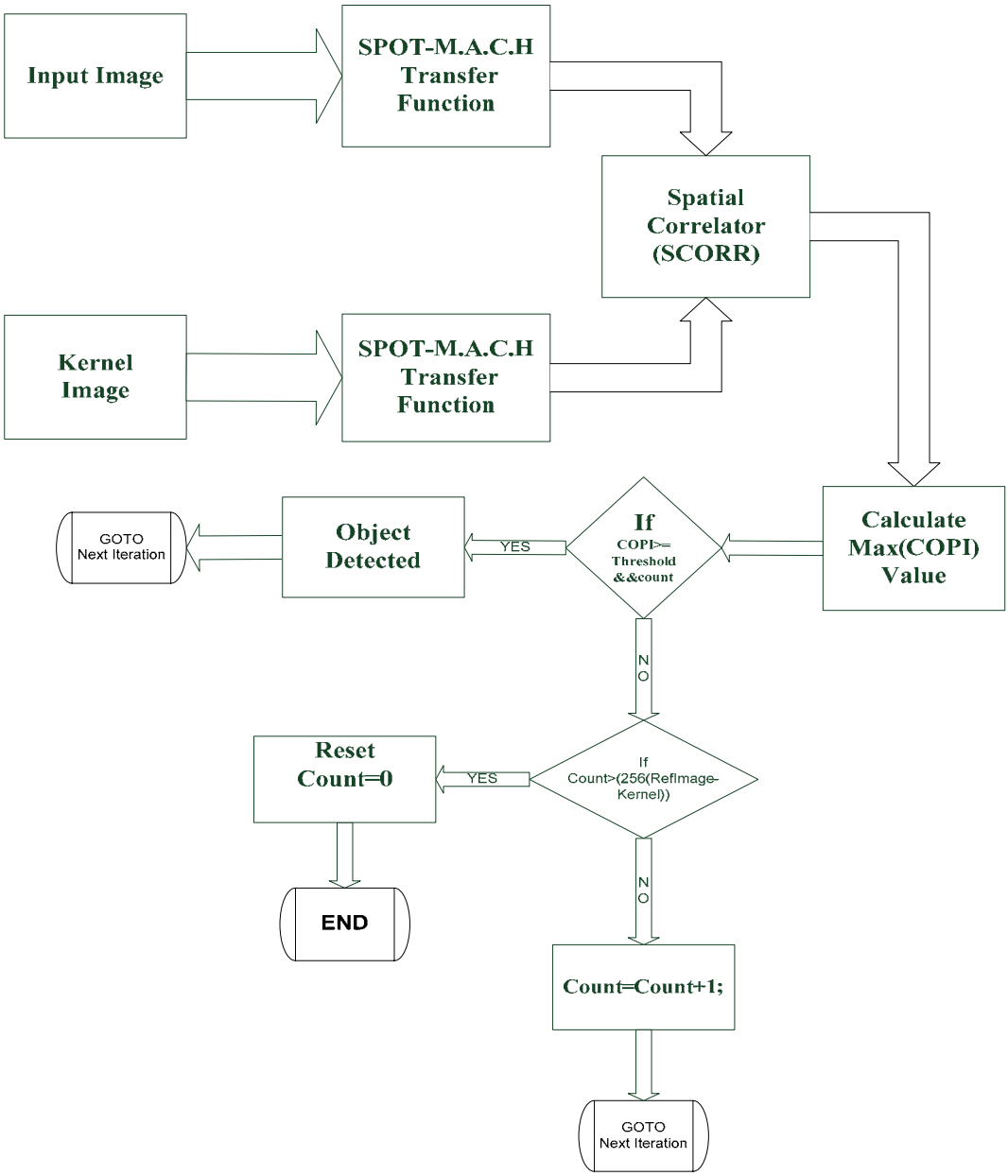


Figure 4.7 The SPOT-MACH Filter

The design and implementation of the SPOT-MACH filter transfer function and its spatial correlator given in Figure 4.7 are further elaborated in the forth coming sections of this chapter.

4.3.1 SPOT-MACH transfer function

In order to formulate the transfer function for the SPOT-MACH filter, first the OT-MACH transfer function should be computed in frequency domain as given by equation 3.39 in Chapter 3.

Considering the above mentioned factors, the SPOT-MACH filter is implemented in the spatial domain using a moving kernel to detect an object in a cluttered background plane of size, typically, between 512x512 and 1024x1024 pixels. To implement the kernel, the filter is designed in the frequency domain using the following adaptation of Equation 3.39 in Chapter 3:

$$h = \frac{m_x^*}{\sqrt{\alpha C + \beta D_x + \gamma S_x}} \tag{4.1}$$

Whereas discussed in Chapter 3 the α, β and γ are non-negative OT parameters, m_x is the average of the training image vector $x_1, x_2, ..., x_N$ (in the frequency domain), and C is the maximum diagonal power spectral density matrix of additive input noise given by $C = max(D_x)$. Normally a white noise covariance matrix is used as additive noise but in this case in order to counter the varying illumination changes, the maximum power spectral density of the input image is added as noise.

D_x is the diagonal matrix with the average power spectral density of the training images as given by equation 3.36 in Chapter 3 , S_x is the similarity matrix as given earlier by equation 3.10.

The reference image spectrum is inverse Fourier transformed to yield a space domain image as can be seen from Equation 4.2.

$$h^* = IFFT\left[\frac{m_x^*}{\sqrt{\alpha C + \beta D_x + \gamma S_x}}\right] \tag{4.2}$$

This is truncated to the target object size which is, typically, less than 256x256 pixels. In addition, care must be taken to store the image as a bi-polar array (rather than an intensity image). The square root operation in the denominator of equation 4.1 is required because the filtering operation must also be applied to the input image

separately so this is pre-processed with the SPOT-MACH transfer function prior to its spatial domain correlation with the SPOT-MACH kernel function. The use of a moving the kernel enables the filter to concentrate on smaller areas of the image and hence minimize the impact of varying illumination patterns. Once the images have been passed through a SPOT-Transfer function they are correlated using the Spatial Correlator which has been discussed in the next section.

4.3.2 Implementation of Spatial Correlator

In the last section SPOT-MACH transfer function was discussed and how it can be used to edge enhance the target image and create composite image from the training dataset. The next step after the training image and the target image have been processed via the transfer function is to perform normalized cross correlation in spatial domain instead of conjugated multiplication in frequency domain.

In this approach a windowing kernel normally the size of the target object is extracted from the composite image created by the SPOT-MACH and is scanned through the target image pixel by pixel. This enables adaptation of the filter dependant on background illumination variances and also enables the normalisation of the filter energy levels. The kernel can be normalized to remove a non-uniform brightness distribution if this occurs in different regions of the image. The main constraint in this implementation is the dependence on computational ability of the system.

As discussed above the main advantage of the correlation function is that it can be normalized for amplitude changes using the correlation coefficient in spatial domain. But the application of normalisation for changes in size and rotation is quite a computationally intensive task. As discussed in Chapter 1 that this can be overcome by using a correlation coefficient when correlating two images. The correlation coefficient can be seen from equation 4.4 given below [10] [11].

$$\varsigma(x,y)=\frac{\sum_s\sum_t\left[f(s,t)-\overline{f}(s,t)\right]\left[w(x+s,y+t)-\overline{w}\right]}{\left\{\sum_s\sum_t\left[f(s,t)-\overline{f}(s,t)\right]^2\sum_s\sum_t\left[w(x+s,y+t)-\overline{w}\right]^2\right\}^{\frac{1}{2}}} \tag{4.3}$$

From Chapter 1 it has been given that $x=0,1,2,.....,M-1$ and $y=0,1,2,.....,N-1$, also $\overline{w}$ is the average value of the pixels in w which is evaluated at the beginning with $\overline{f}$ being the average value of f in the region which contains the moving window. The summations are computed in the regions which are overlapping and have common coordinates for f and w. The coefficient of correlation $\varsigma(x,y)$ is scaled between the range of -1 to 1 which is independent to the scale changes in the amplitude of f and w [1] , [2] and [11].

The reason for applying normalisation is due to the failures of OT-MACH in the frequency domain by not being able to cope with illumination variations and hence giving false detections. The frequency domain approaches are un-normalized in nature and this proves that when using a fixed sized window depending on the size of the target object using equation 4.1 gives a normalized approach for correlation which enables pattern matching even in the areas of non-uniform brightness.

As discussed earlier in Chapter 1 that when the object is oriented at an unknown angle than $w(x,y)$ must be rotated so that it aligns with the degree of rotation of $f(x,y)$ which is quite an exhaustive process. Due to the impracticalities involved in this process correlation is seldom used in scenarios where unconstrained rotation is present in the target scene [11] [17]. But this has been overcome by the SPOT-MACH transfer function which creates a spatial domain composite image of multiple orientations that can be effectively used as a kernel to perform the normalized correlation with the target scene.

The local window normalisation offers advantages in situations where the global energy for the scene is nearly equal for each orientation of the target in this case applying global normalisation would have a minimal effect. Whereas in the case of using local window energy normalisation each orientation of the target object would be normalized relatively independently.

The designed Spatial Correlator (SCORR) for the scope of this thesis has the ability to extract a window from the composite image and then effectively scan it through the target image pixel by pixel. The working of the SCORR has been explained with the help of the following experiment where the capabilities of the SCORR have been tested directly without using the SPOT-MACH transfer function.

In this case a 512x512 image containing a brightly illuminated square having pixel intensity of 250 similar to the one used for the OT-MACH tests in frequency domain has been used. A 64x64 window containing a brightly illuminated square has been extracted as the moving window and SCORR has been used to perform the correlation which can be seen from Figure 4.8 below.

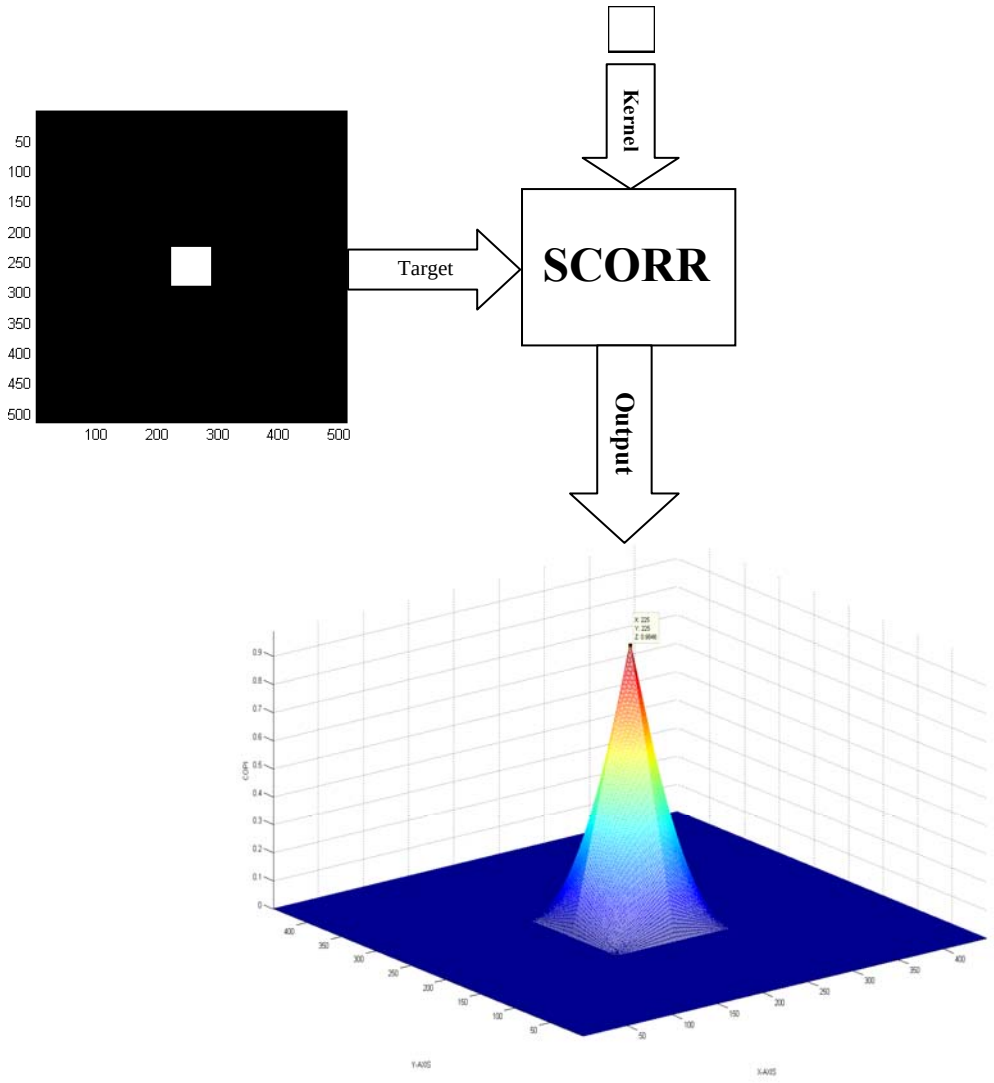


Figure 4.8 Working of Spatial Correlator using auto-correlation

From Figure 4.8 it can be seen that when a target image and kernel window, containing a square, are correlated together using SCORR then a pyramid shaped correlation output plane is created which is normalized between 0 and 1 having the maximum COPI value of 0.98 which signifies that the SCORR is working accordingly. The only thing to be careful of when using the SCORR is the kernel size as choosing the wrong kernel size will not give an ideal correlation output plane. If the target object is not covered completely by the window than a scaling loss will occur. In the case of scaling loss, the correlation output plane achieved can be seen from Figure 4.9 below.

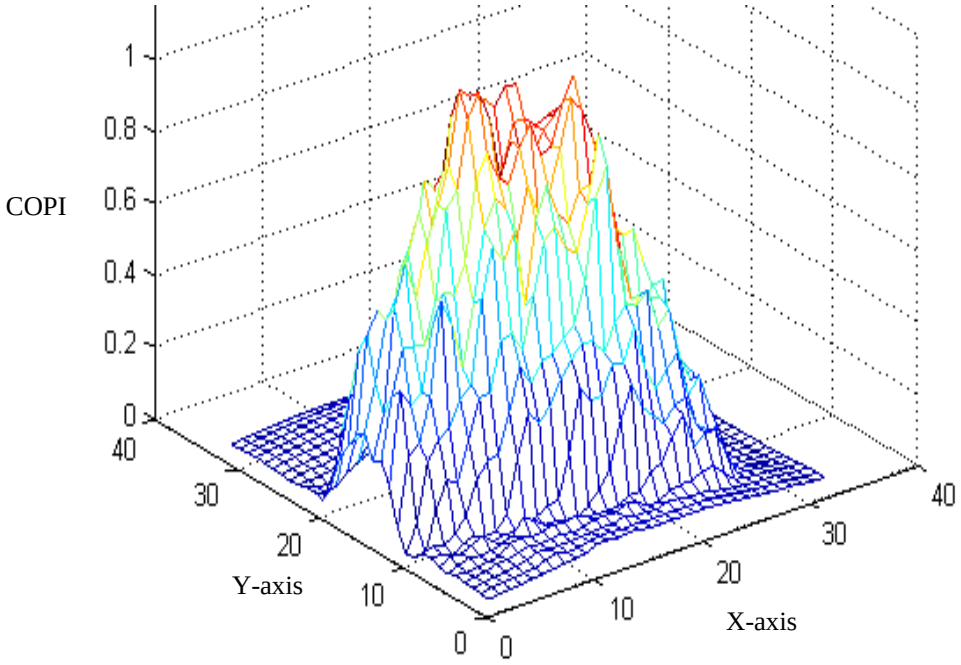


Figure 4.9 Scaling loss effects in the SCORR

From Figure 4.9 shows the correlation plane when a 16x16 window is chosen to detect a square of size 24x24 in a 48x48 image, the following losses occur. It can be seen that multiple peaks are achieved and the shape of the pyramid is not well defined as compared to Figure 4.8. This signifies that the kernel was unable to find an exact match of the square. This shows that the window size is an integral part in the operation of SCORR.

Once the working of the SCORR was established, the SPOT-MACH transfer function can be applied to the images to obtain the correlation output plane in the spatial domain and the target object can be thus detected and located.

4.3.3 Working of SPOT-MACH filter

In this section images are now passed through the SPOT-MACH transfer function and the SCORR and a correlation output plane is produced. By applying the algorithm shownin Figure 4.7 and using the same set of images shown in Figure 4.8, the following results have been achieved.

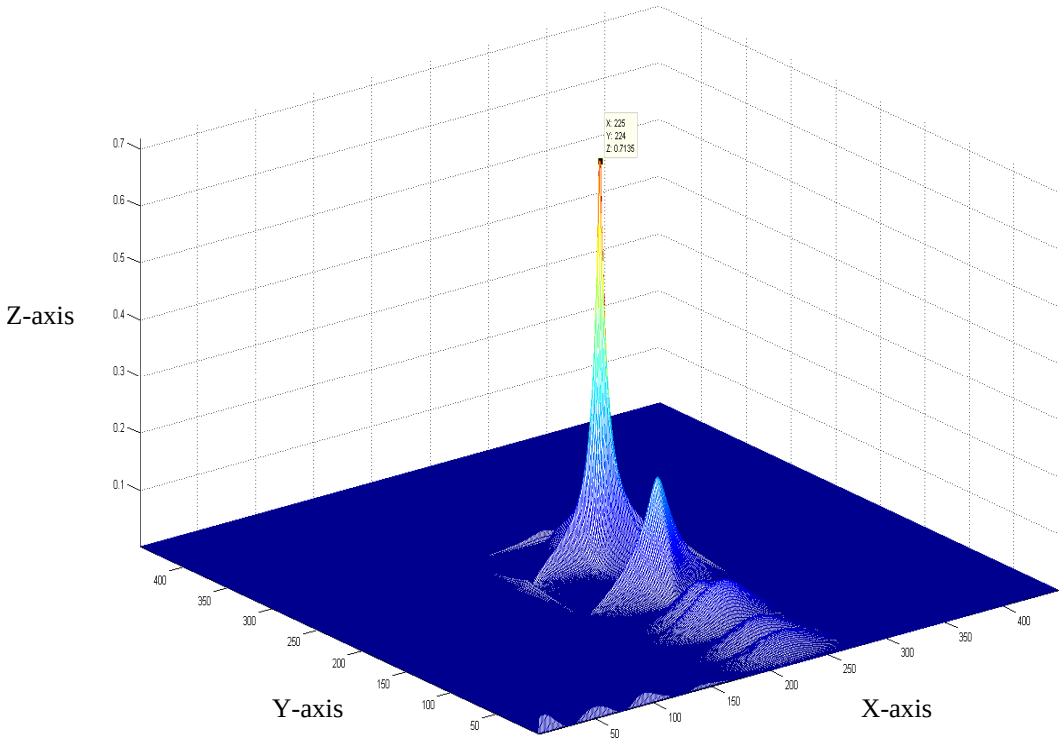


Figure 4.10 Correlation Output Plane using SPOT-MACH filter

From Figure 4.10 it can be seen that with the application of the SPOT-MACH filter the correlation output plane has changed as compared to the SCORR output shown in Figure 4.8. A sharp peak can be seen in the centre which gradually reduces down due to the movement of the kernel across the image. The SPOT-MACH transfer function has been tuned using parameter values $\alpha = 0.001$, $\beta = 0.7$ and $\gamma = 0.01$. The pyramid shaped

plane has been transformed into a peak having a $COPI$ of 0.719 which signifies a good detection.

A further point to be noted in the case of the SPOT-MACH filter is the size of the correlation output plane which is smaller as compared to the frequency domain plane. This is due to the fact that the windowing function is used which will effectively return a lesser number of correlation points if zero padding is not used. The correct size of the correlation plane can be established using the equation 4.4.

$$Corr\_plane\_size(x, y) = (x - window\_size, y - window\_size) \quad (4.4)$$

Since the correlation output plane size has been changed and the location of the peak corresponds to the location of the object. Thus in order to determine the corrected location of the target object the following transformation should be applied to the peak coordinates of the SPOT-MACH filter

$$Corr\_Outputplane = \left(x - \frac{winsize}{2}, y - \frac{winsize}{2} \right) \quad (4.5)$$

The use of equation 4.5 will now give the exact location of the object in the correlation output plane for the SPOT-MACH filter.

Having established that the SPOT-MACH filter is able to detect similar shaped objects, its capabilities have been put to test using the same dataset employed to in section 4.1 for the OT-MACH filter in frequency domain. The main aim of this test is to assess the tolerance of the SPOT-MACH filter to the change in illumination when a false target is present in a brightly illuminated area. In order to correctly compare the difference in performance for the frequency domain OT-MACH filter and the SPOT-MACH filter, a similar set of reference and target images will be used.

In this case, the target image has been taken as that shown in Figure 4.1 and the reference image as in Figure 4.2, i.e. similar to the images used for the OT-MACH frequency domain filter assessment. The SPOT-MACH correlation output plane is shown in Figure 4.11 below.

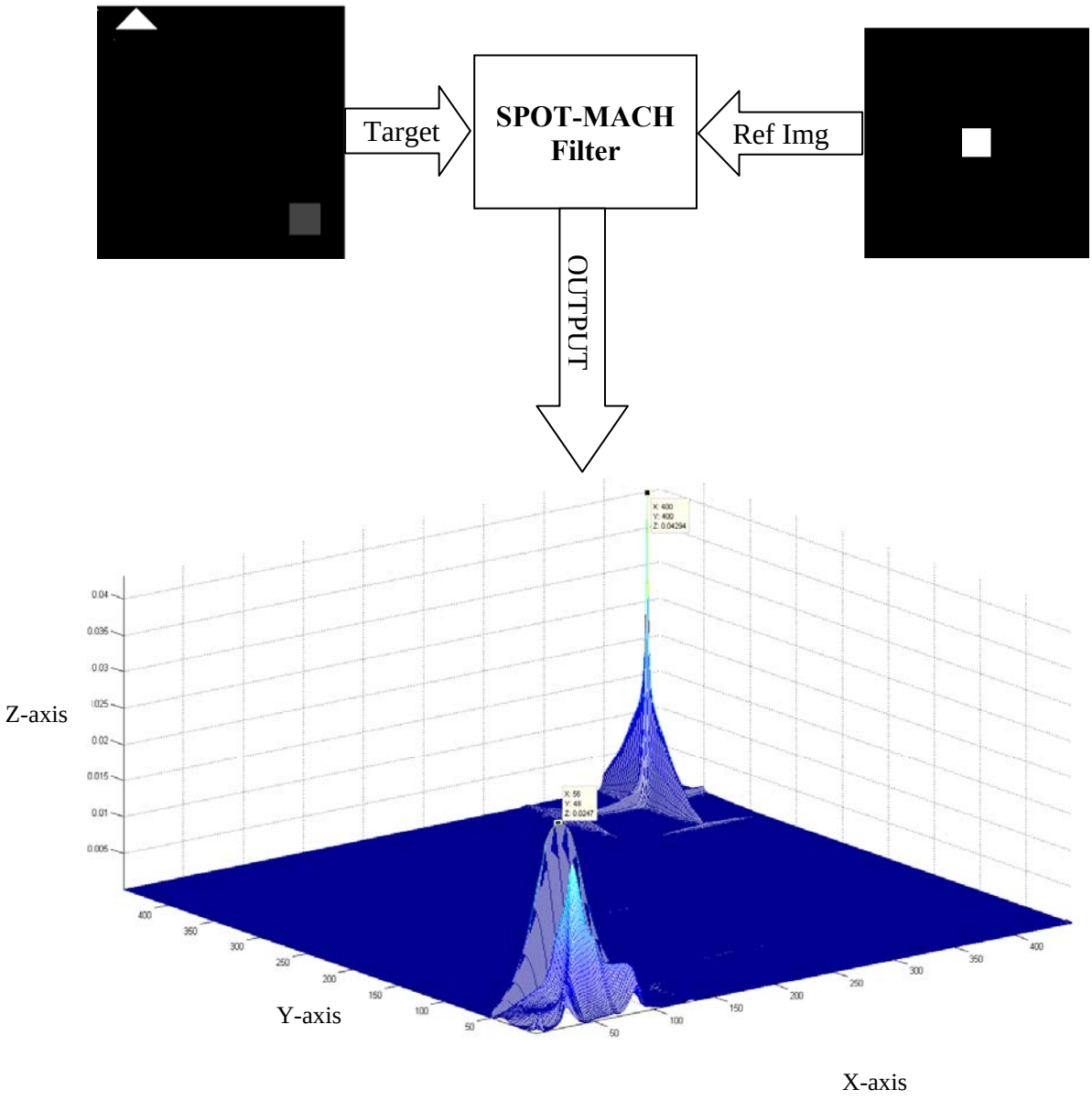


Figure 4.11 SPOT-MACH correlation output plane for illuminated false target

From Figure 4.11 it can be seen that the $COPI$ for the false target is 0.024 and for the true target is 0.0429.As compared to the results for the frequency domain OT-MACH, which was unable to detect the true target in the presence of a brightly illuminated false target, the SPOT-MACH has been successful.

It can be deduced from this result that even with the improved discrimination capabilities the frequency domain OT-MACH, it was unable to distinguish between true and false targets, as shown in Figure 4.4. The SPOT-MACH achieves this by applying local window energy normalization which effectively suppresses the high energy values of the false target and normalizes with respect to neighbouring pixel intensities.

From figure 4.11 it can be seen that the SPOT-MACH is tolerant to changes in illumination and would be effective in situations where the object to be detected might be present in a shaded area. Another major advantage of the SPOT-MACH filter is that the quality of detection is easily verifiable in terms of the highest peak which corresponds to the detected target with its coordinates (after being translated using equation 4.6) corresponding to the exact target location. In contrast, since the frequency domain implementation is unconstrained it is difficult to assess what signifies a good detection other measures such as PCE and PSR have to be considered to establish a criteria for detection. This makes SPOT-MACH an easier approach when determining the quality of detection.

In the next section the performance of the SPOT-MACH is assessed based on experimentation with a visible imagery data set containing a car oriented at multiple angles and an image containing background clutter.

4.4 Performance analysis of SPOT-MACH

In this section the performance of the SPOT-MACH filter has been tested using the visible imagery data set of a car which has been employed previously in Chapter 2 and 3 for testing the capabilities of the SDF and frequency domain OT-MACH. Also, a new data set has been used which is a 15% scaled version of the original car dataset.

The experiments have been conducted using a visible imagery scene where there is a uniform distribution of energy across the image plane. A target image with a uniform distribution of energy has been used in order to compare the performance of the SPOT-MACH with the SDF and frequency domain OT-MACH filter.

The scaled version of the training image data set has been used to assess the effect of scaling on the performance of the SPOT-MACH filter. Being a windowing based approach; the use of a scaled data set could degrade the performance of the SPOT-MACH.

For the case of testing the tolerance to change in illumination of the SPOT-MACH filter, a new dataset has been created using the car images at different orientations but with a varying set of background images containing non-uniform distributions of energy in order to assess the filter performance in variable lighting situations. These experiments are discussed in more detail below.

4.4.1 Experiments with Uniform Brightness Background

In order to determine the capabilities of the SPOT-MACH filter as compared to the frequency domain OT-MACH a set of tests were conducted using the same visible imagery dataset oriented from 0 to 360 degrees as used in Chapter 3.

The car image training dataset used for the SPOT-MACH filter is the same as given in Chapter 2 containing car images oriented between 0-360 degrees with a 10 degree step.

The target image used for this experiment contains a car oriented at ten degrees containing background clutter. This is shown by Figure 4.12 given below.

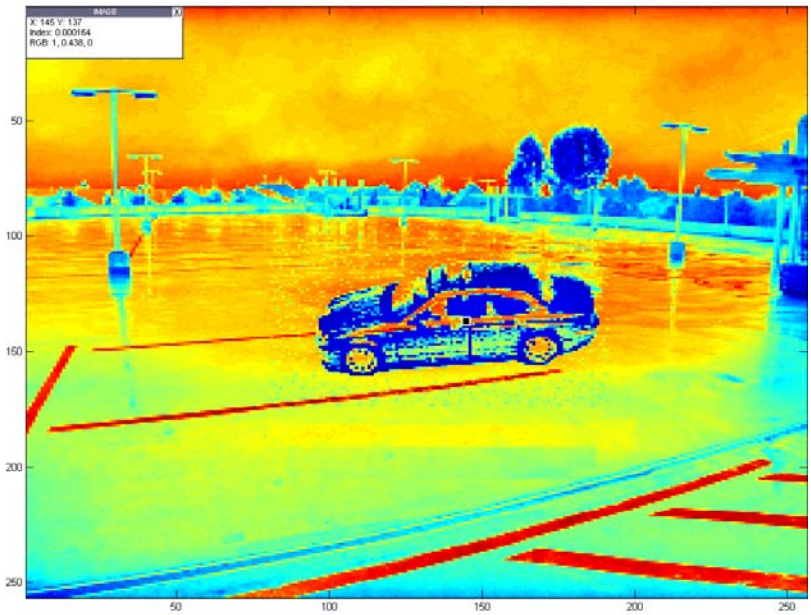


Figure 4.12 Target image for SPOT-MACH

As can be seen from Figure 4.12 the energy distribution is uniformly spread out over the whole target scene. The location of the target is at $(145,137)$ as can be seen from the small black dot on the target.

When the SPOT-MACH filter was applied to the target scene shown in Fig 4.12 keeping the parameters the same as OT-MACH the frequency domain OT-MACH filter i.e. $\alpha = 0.001$, $\beta = 0.7$ and $\gamma = 0.01$ the following set of results were obtained.

Initially the test had been performed by assessing the auto correlation values for the same image, this being a car oriented at zero degrees. These results are shown in Figure 4.13.

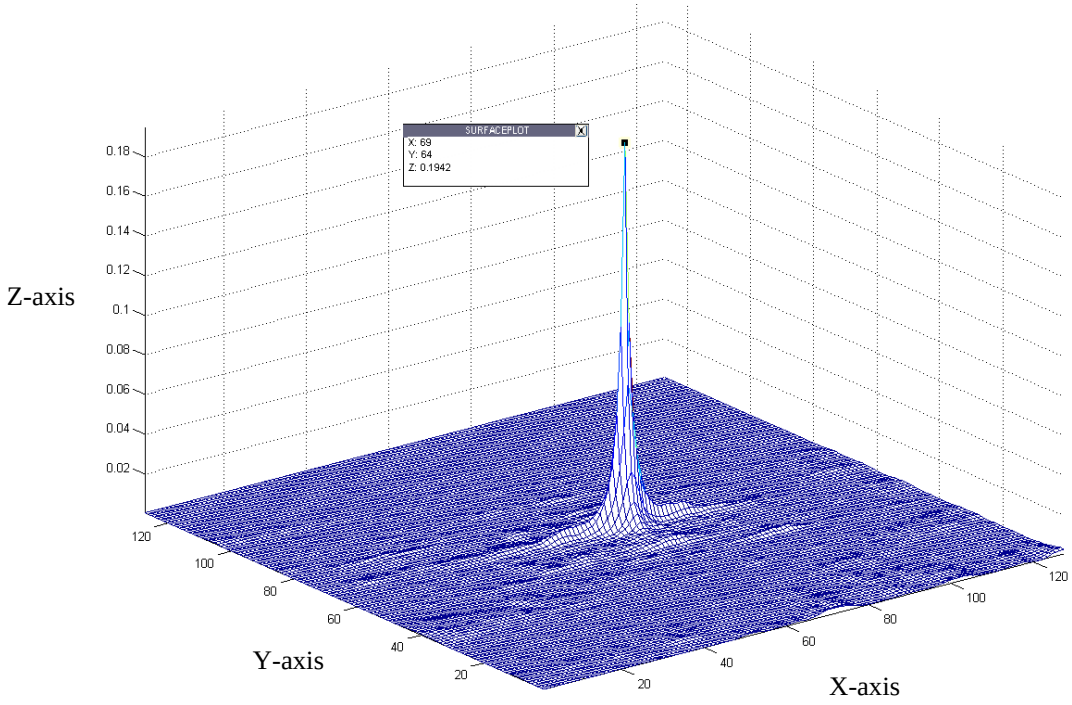


Figure 4.13 SPOT-MACH Autocorrelation plane for Car at 0 degrees

From Figure 4.13 it can be seen that the *COPI* for the autocorrelation is 0.1942, and in this case no background clutter is present in the target scene. In order to determine the capabilities of the SPOT-MACH in terms of discrimination of the target in the presence of background noise, a set of tests were conducted with the target and a filter trained on orientations from 0 degrees to40 degrees, as shown in Figure 4.14.

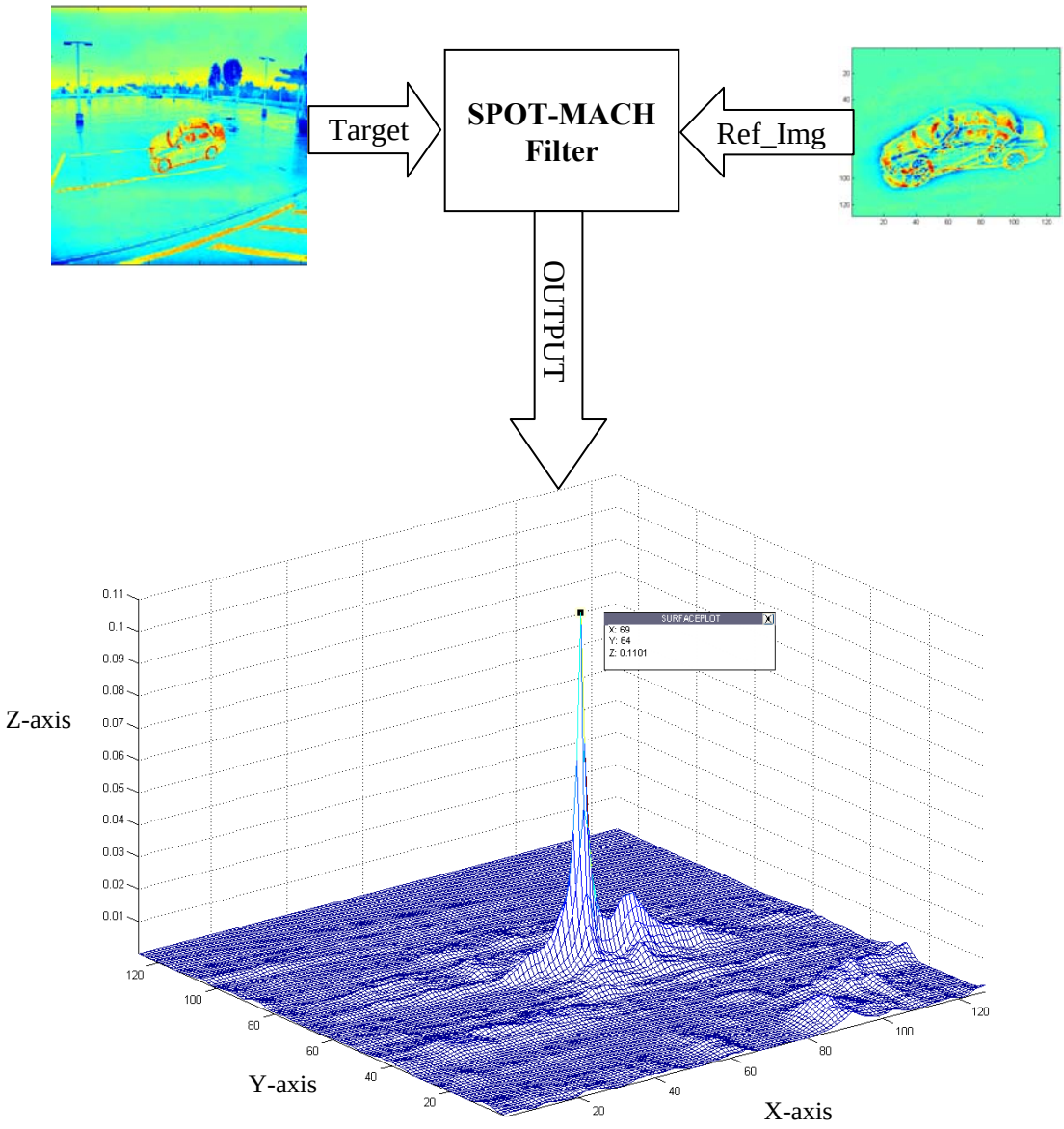


Figure 4.14 SPOT-MACH Correlation output plane for car reference filter oriented between 0 to 40 degrees

From Figure 4.14 it can be seen that when a 256x256 input scene containing background clutter is filtered using a 128x128 window containing a composite image of the car oriented from 0 to 40 degrees, the correlation output plane contains a peak having a *COPI* value of 0.1101. From the correlation output plane it can be seen the target object was successfully detected at location (139,129) which is the edge of the desired target object.

4.4.2 Experiments with 15% scaled training image dataset

In this section the performance of the SPOT-MACH has been tested by training it with a 15% size reduced dataset. This is due to the fact when the viewing distance of the target changes, the target may appear smaller in size. In this case the window size should be changed accordingly as well, as otherwise a scaling loss will occur.

When correlation is performed with the same parameters as used in the previous experiments, the following results are obtained, as shown in Figure 4.16.

Figure 4.15 15% scaled car image database

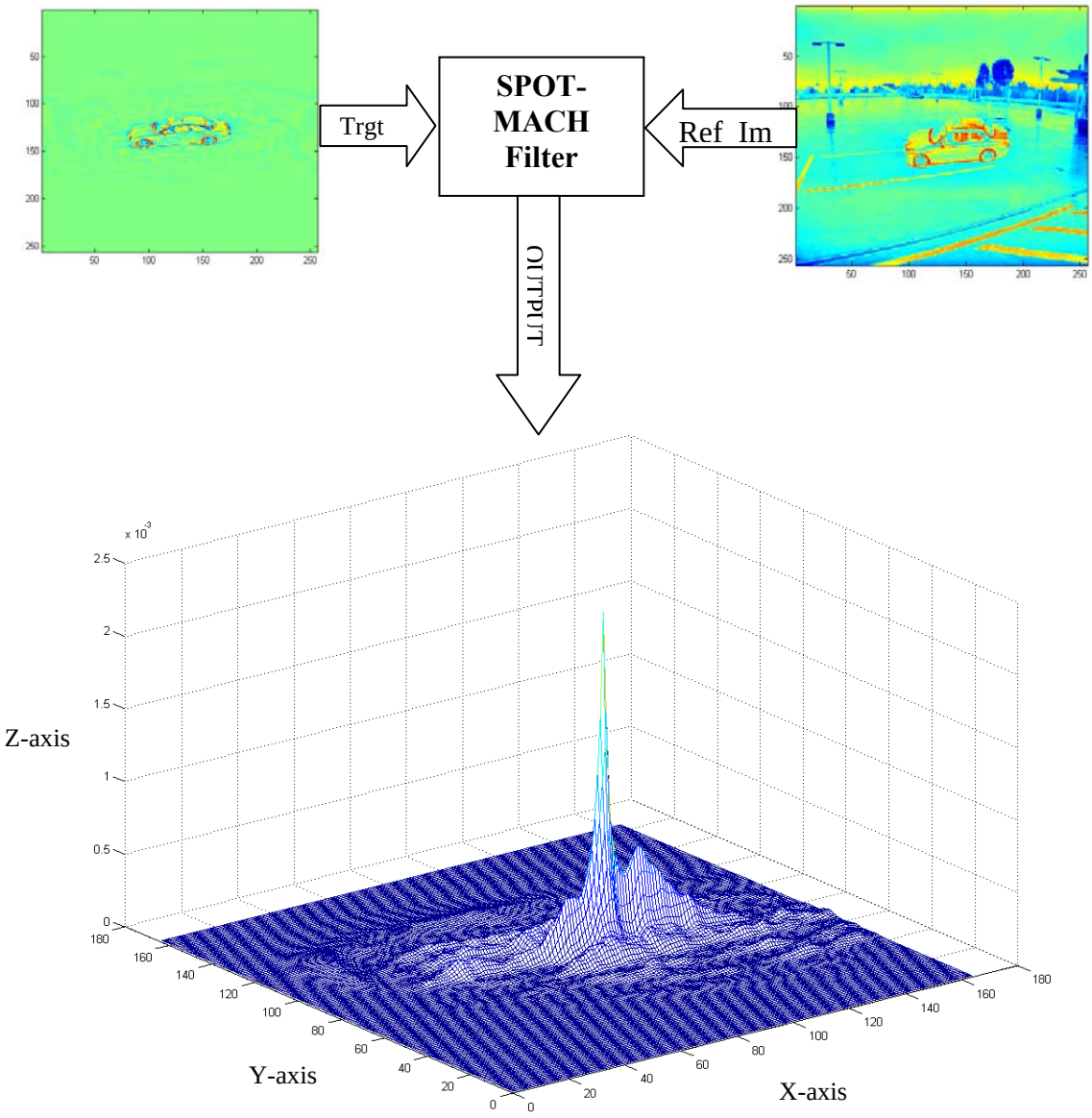


Figure 4.16 Spatial Correlation with change in scale

As can be seen from Figure 4.16 by changing the scale and keeping the parameters the same the target has been detected but the side lobes and background noise are increased. The *COPI* in this case is given by 2.5×10^{-3} which is extremely low due to the fact that scaling of the training image dataset has reduced the size of the object within the windowing kernel and an effective correlation cannot be attained. This shows that there is degradation in performance of the SPOT-MACH filter in the presence of a scaling coefficient.

4.4.3 Experiments with non-uniform brightness dataset

It has been seen from section 4.4.2 that the performance of the SPOT-MACH is degraded when there is a scaling affect present in the training image dataset. However, the main advantage of the SPOT-MACH filter is its tolerance to change in illumination when the target object is present in a shaded area and there is non-uniform distribution of energy across the target scene, a situation in which frequency domain approaches fail.

In order to effectively test the capabilities of the SPOT-MACH a new dataset is created which contains the car oriented at 0, 90, 180 and 360 degrees. A set of scenes have been created where the lighting effects have been varied to create non-illumination patterns with certain areas having focussed illumination on false targets. The car image dataset is shown in Figure 4.17 below.

Figure 4.17 Car image dataset for illumination tolerance testing of the SPOT-MACH filter

In section 4.1 it was shown that the SPOT-MACH filter is tolerant to change in the illumination and it tends to detect the target where the frequency domain OT-MACH fails. In order to assess the performance of the SPOT-MACH in a real-life scenario, a dataset was created using visible imagery which shows target objects at different orientations and with varying illumination patterns. A more detailed comparison of the SPOT-MACH and frequency domain OT-MACH is conducted in Chapter 6. In this

section, a target image has been selected which has an object in a dark shaded region and poses a very demanding testing situation for the correlation filter. The target image employed is shown in Figure 4.18 below.

Figure 4.18 Target Image containing non-uniform illumination and dark region

The target image given in Figure 4.18 was correlated using OT-MACH using the frequency domain OT-MACH filter trained to detect a single target oriented at 0 degrees using the training image dataset shown in Figure 4.17 and with the parameters where $\alpha = 0.001$, $\beta = 0.7$ and $\gamma = 0.01$. The correlation output plane is shown inFigure 4.19 below.

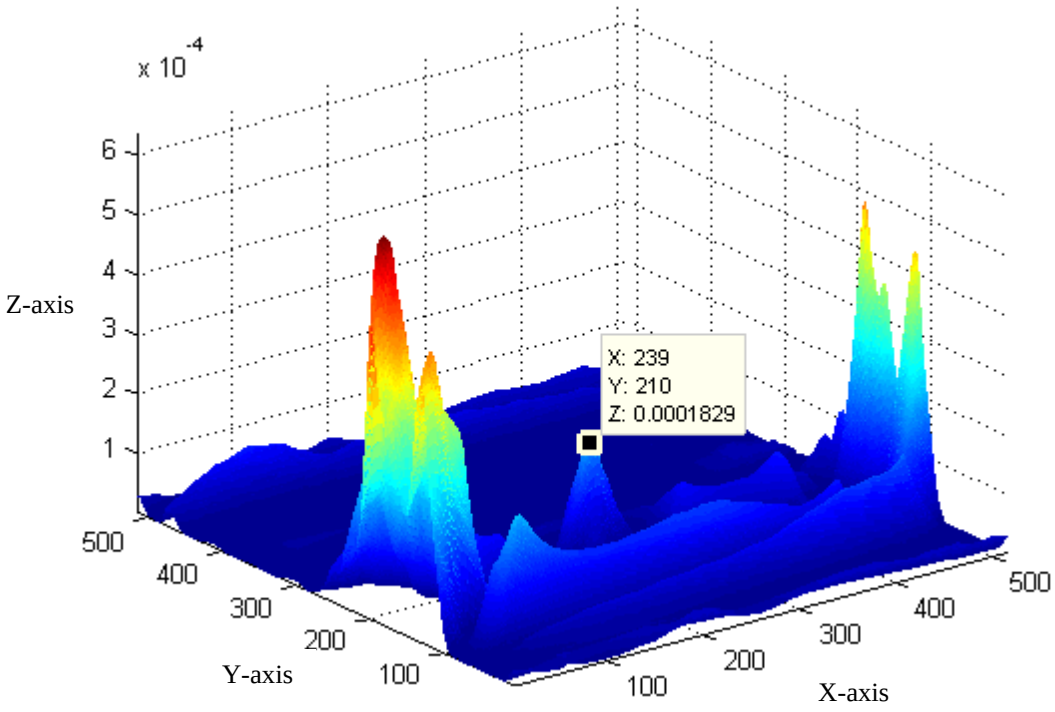


Figure 4.19 Correlation Output Plane for frequency domain OT-MACH filter

It can be seen from Figure 4.19 that the frequency domain OT-MACH filter fails to detect the object which is located at the centre of the target image. In Figure 4.18 it can be seen that there is a lot of energy distribution where the light is shining in the top-right corner of the target image. From the correlation output plane shown in Figure 4.19 it can be verified that the frequency domain OT-MACH filter gives false detections if a certain region is more strongly illuminated which can be seen from the large peaks in the correlation plane where illumination is present. The *COPI* value is 0.0001 which is extremely low as compared to the other peaks in the correlation plane which shows that the frequency domain OT-MACH filter has been unable to detect the object in the shadow region of the image.

In order to assess the performance of the SPOT-MACH filter, the same dataset was used and a correlation output plane was generated as shown in Figure 4.20.

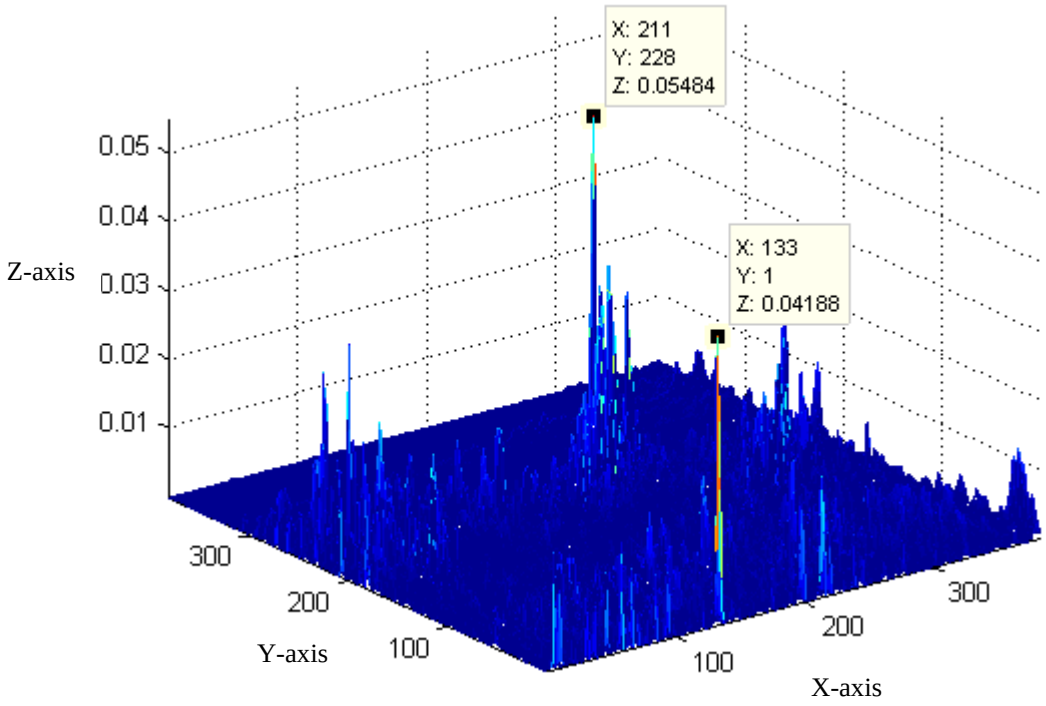


Figure 4.20 Correlation Output Plane for SPOT-MACH filter

From Figure 4.20 it can be seen that in the case of the SPOT-MACH filter using a local window of 128x128 within a scanned target image of 512x512, a peak at the centre of the scene has been produced. This proves the point that the use of local window normalisation overcomes the effect of non-uniform energy distribution and the object can be successfully detected.

Clearly it has been seen that the SPOT-MACH filter provides an advantage over its frequency domain implementations, with the only problem with the SPOT-MACH filter being the computational resources required when compared to the frequency domain OT-MACH filter.. In the case of real time processing, a delay in the detection of the target object is not acceptable and hence a set of enhancement techniques have been proposed to reduce the detection time when using the SPOT-MACH filter and make this comparable to the frequency domain OT-MACH filter. The proposed techniques are discussed in more detail in Chapter 5 of this thesis.

4.5 Conclusion

In this chapter a space variant approach to implement a correlation filter has been discussed. One of the main advantages in the spatial domain is the ability to allow localized normalization of the filter which is not possible in the frequency domain. But normalization is a time consuming process to implement in the spatial domain. The proposed SPOT- MACH filter has been created using a training image data set of car images using visible imagery.

In order to search for the desired target in the scene, the SPOT-MACH filter searches the entire target image step by step and the correlation process has to be repeated for each location in the image. To overcome this issue, a robust moving window mechanism has been proposed in this chapter. The system takes a sub-image from the input image and its size is made equal to the window which in turn is user dependent.

The size of the window should be small enough to reduce the simulation time and large enough to cover the distortions in the target object. In this chapter the SPOT-MACH filter has been correlated with a sub-image. A threshold was calculated by correlating each image with the MACH filter and taking the average of their resultant correlation peak intensities. Correlation peak values above the threshold imply that the target is an in-class object. All the out-of-class objects will have values lower than the threshold.

Also in this chapter a comparison has been performed with the OT-MACH filter implemented in the frequency domain and it has been shown that the performance of the OT-MACH is degraded in the presence of non-uniform illumination patterns within a target image. The SPOT-MACH has been shown to overcome these problems by using local window energy normalization but with the caveat that the size of the window greatly influences the performance of the SPOT-MACH.

The only advantage the OT-MACH frequency domain implementation is the computational speed. The time of detection when employing the SPOT-MACH form a major difficulty as this tends to prolong the processing time for the filter excessively. In order to overcome these issues a set of enhancement techniques has been proposed which will be discussed in more detail in Chapter 5, effectively making the SPOT-

MACH a two stage detection filter where the first stage is optimization and the second stage detection.

Chapter 5

Speed Enhancement Techniques for the SPOT-MACH Filter

5.1 Introduction

In Chapters 2 and 3 a brief discussion was given on the history of correlation filters and the formulation of the unconstrained correlation criterion which led to the development of the MACH filter and later to its enhancement the OT-MACH filter. In Chapter 4 a moving window was used to implement a Spatial Domain equivalent of the OT-MACH filter which was named the SPOT-MACH filter which had the ability to be locally modified depending upon its position in the input frame. This enabled the adaptation of the filter dependant on locally variant background clutter conditions and also allowed the normalisation of the filter energy levels at each step.

The SPOT-MACH filter was shown to offer an advantage over its frequency domain implementation as shift invariance was not imposed on it. Also it was shown that the performance of the frequency domain OT-MACH filter becomes degraded in the presence of non-uniform illumination patterns within a target image. The SPOT-MACH filter was shown to overcome the problem of illumination changes by using local window energy normalisation.

The only drawback associated with the spatial domain approach was that the window size is directly proportional to the required computational resources and would adversely affect the performance of the system if the window size is not proportionate to these resources. This makes the SPOT-MACH filter a very computationally expensive filter in terms of the complexity of computations. In the case where a 64x64 window is used to move across the target image at every window shift there are 4096 multiply and accumulate functions that have to be performed for every shift of the window over the

size of the image. If the window size is increased to 128x128 to scan an image of total resolution 512x512 pixels then a total of 16384 executions of the correlation function, given by equation 1.4, have to be performed 147456 times; hence the computational complexity of the spatial domain approach. In order to make the SPOT-MACH filter comparable to its frequency domain counterparts a set of speed enhancement approaches have been proposed, effectively making it a two stage detection filter. These techniques are described in this Chapter.

The SPOT-MACH filter is computationally intensive but tolerant to changes in illumination. In this section, the enhancement capabilities of the SPOT-MACH filter are explored and potential application areas have been identified. The optimisation of SPOT-MACH filter can take place in both hardware and software domains based on its target application. A total of four enhancement techniques are developed for the SPOT-MACH out of which two are software based and the other two hardware based. The developed techniques for enhancement are:

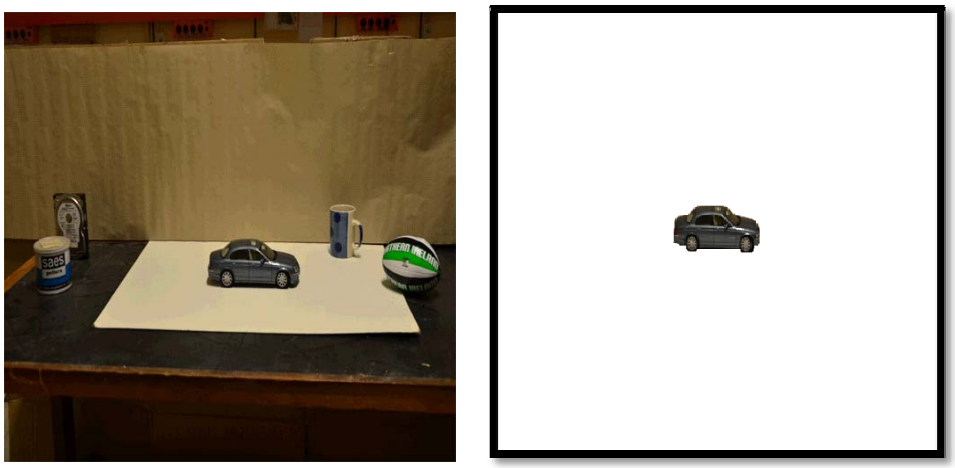
1. Speed Enhancement using Low-Pass Filtering
2. Speed Enhancement using Entropy based filter
3. GPU based optimisation technique
4. FPGA based optimisation technique

These software and hardware based techniques are discussed in more detail in the forthcoming sections of this chapter.

5.2 Software Based Enhancement Techniques for the SPOT-MACH filter

In order to establish a benchmark for timing analysis of the SPOT-MACH filter a 512x512 image was used which contains a car image under uniform lighting. The choice of testing hardware for the software based techniques was a PC with an INTEL Core 2 Duo 2.10 GHz processor and 4 Giga Bytes of Random Access Memory running a 32 bit operating system.

The training and the reference images for the timing analysis are shown by Figure 5.1 below.



(a) **(b)**

Figure 5.1 Target image for timing analysis (a),
Training image at 10 degree orientation (b)

Initially the timing results using the profiling function in MATLAB are calculated for the frequency domain OT-MACH filter when applied to the image shown in Figure 5.1 which is shown in Figure 5.2.

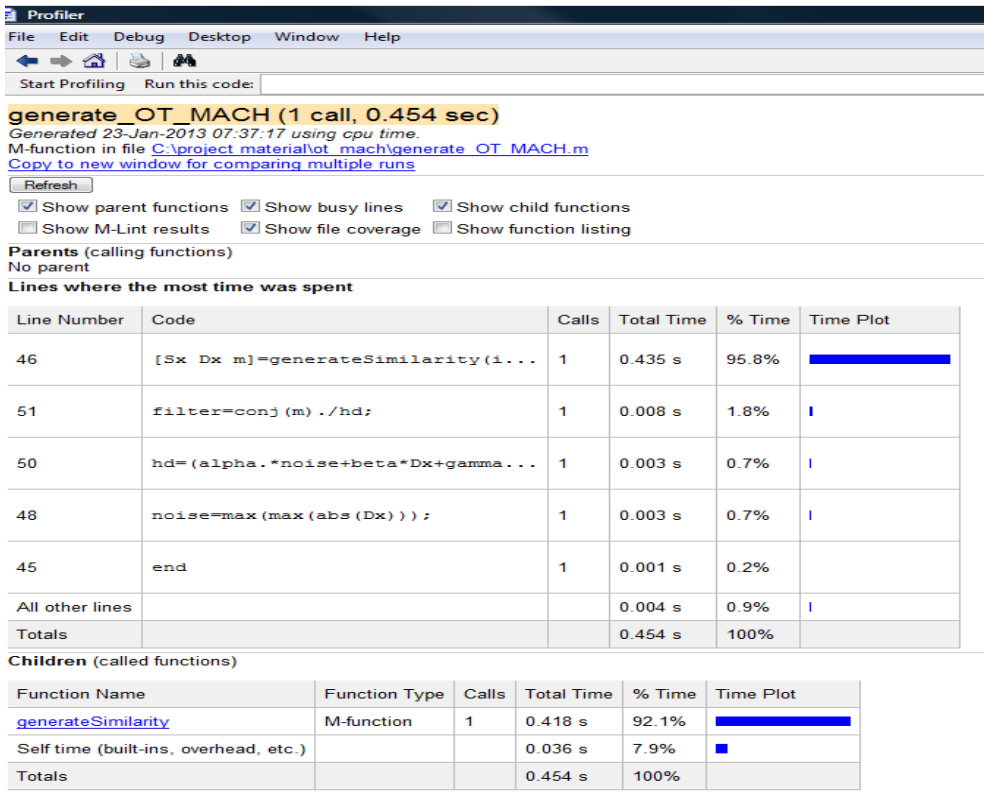


Figure 5.2 Frequency domain OT-MACH filter timing analysis for 512x512 images

From Figure 5.2 it can be seen that the total time for the frequency domain OT-MACH filter to run is $0.436 + 0.435 = 0.871s$, these results can be effectively used to compare the timings for the SPOT-MACH filter. These techniques are discussed in more detail in the forthcoming sections of this chapter.

5.2.1 Speed Enhancement using Low-Pass Filtering

An optimisation technique is proposed in this section for the SPOT-MACH filter which employs an optimisation technique using low-pass filtering to highlight the potential regions of interest in the image and then restrict the movement of the kernel to these regions to allow target identification.

It was established in the Chapter 4 that the SPOT-MACH filter offers the advantage that it does not have shift invariance imposed on it as the kernel can be modified depending upon its position within the input image. This allows normalisation of the kernel and as well as inclusion of a space domain non-linearity to improve performance.

In order to evaluate the detection time of the filter, MALAB profiling functions are used. In this case when the reference and the target images are acquired, the target image can be passed through a low pass filter to highlight potential regions of interest by showing a high energy distribution in areas where any object might be present.

These regions are extracted and the moving window is restricted to locate only in these regions, effectively reducing the number of computations required. The developed algorithm is summarised in Figure 5.3.

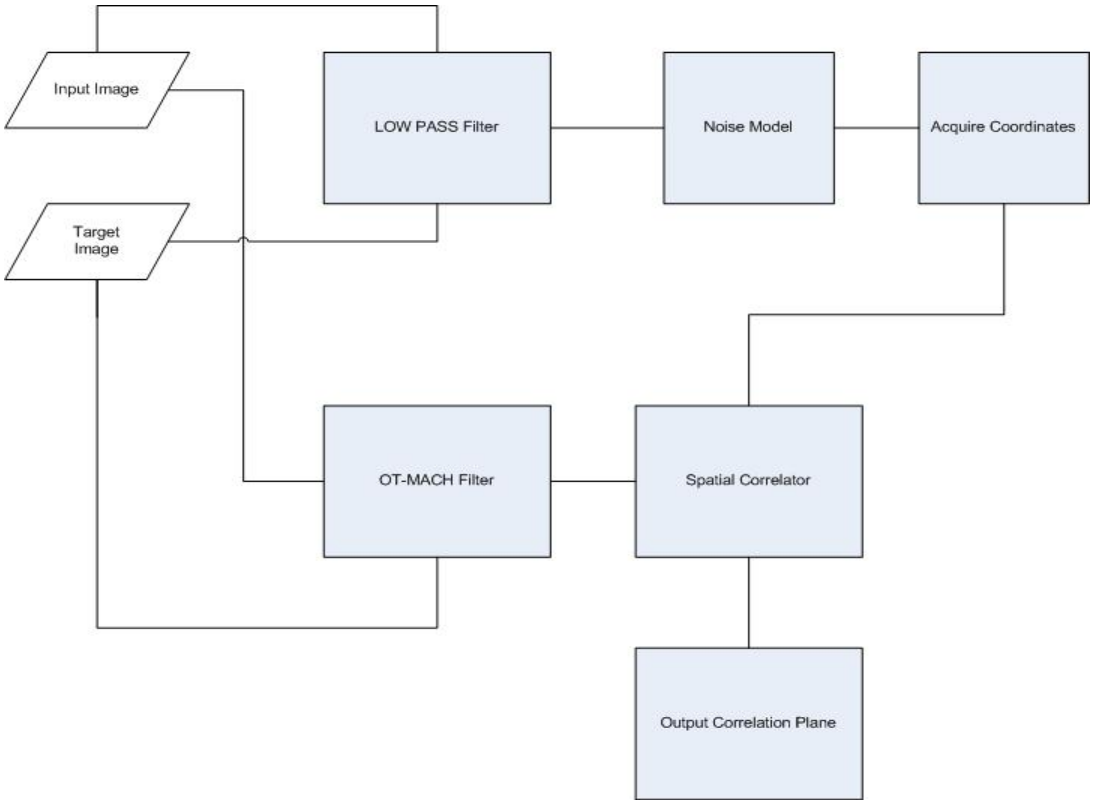


Figure 5.3 Speed Enhancement Algorithm using Low Pass filter for SPOT-MACH

The developed optimisation technique shown in Figure 5.3 involves a low pass filter which can be a frequency domain OT-MACH filter transfer function tuned to behave as a low pass filter. The optimisation technique uses estimation from the filtered values and predicts areas of interest for locating the windowing kernel to minimise the overall processing time but at the same time retaining the detection capabilities of the filter. It is a two stage approach where initially coordinates are calculated for the windowing kernel. In the second stage the OT-MACH filtered images are passed through the spatial correlator where the coordinates acquired in the previous stage define the movement of the kernel. The kernel uses a neighbourhood averaging technique to define a region of interest for the windowing kernel. After scanning the desired region of interest the kernel jumps to another region within the image effectively skipping unwanted area to minimise the computation resources required.

There are a lot of heuristic adaptive-neighbourhood approaches for filtering images corrupted by background noise. Instead of using fixed-size, fixed-shape

neighbourhoods, statistics of the noise and the signal are computed within variable-size, variable-shaped neighbourhoods that are selected for every point of estimation [80].

The technique employed, as summarised in Figure 5.3, when applied to the image shown in Figure 5.1 generates the region of interest map shown in figure 5.4.

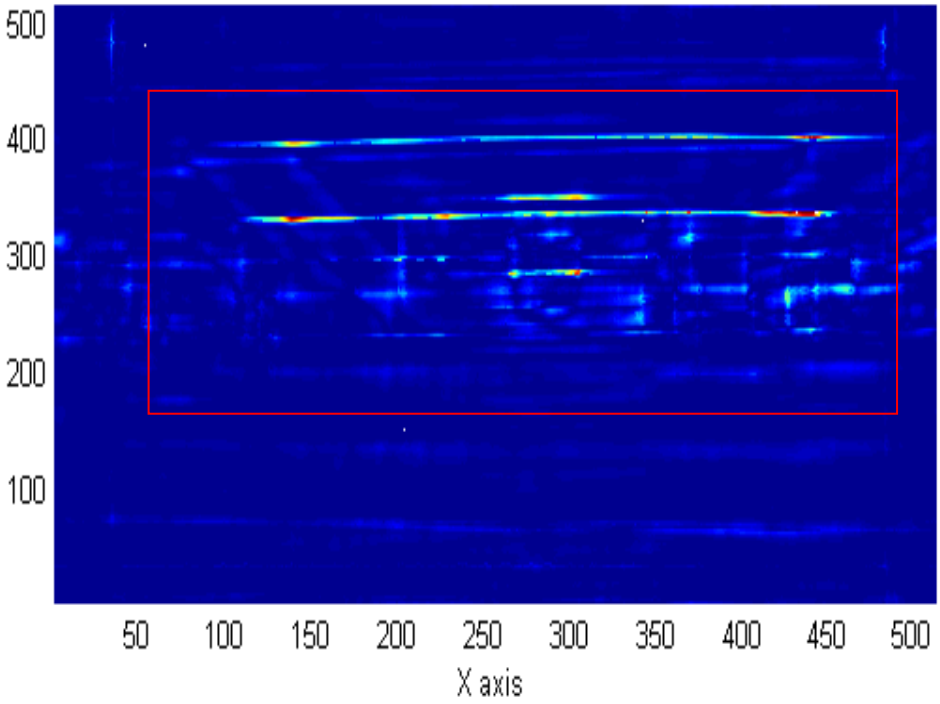
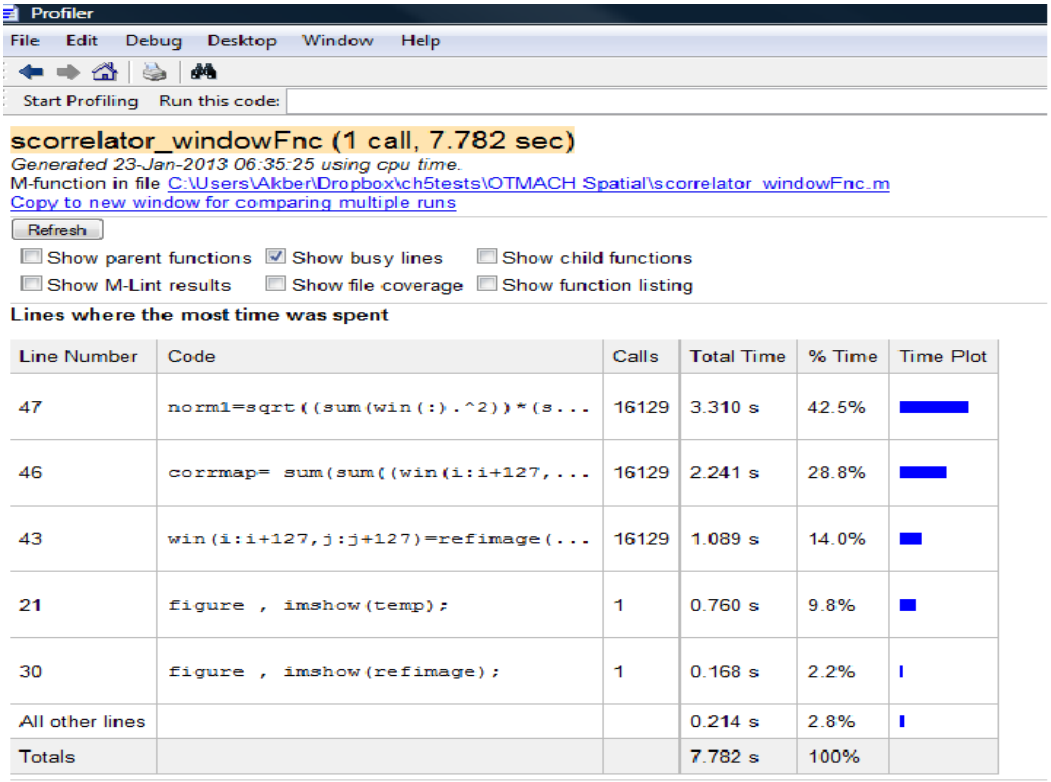


Figure 5.4 Generation of Region of Interest by low-pass filter for the SPOT-MACH filter

In Figure 5.4 the regions of interest for the SPOT-MACH filter is highlighted and in order to evaluate the gain in performance this was compared with the results obtained without enhancement. From Figure 5.4 it is shown that the low-pass filtering has resulted in reducing by almost 40% of the kernel window movement.


Profiler

File Edit Debug Desktop Window Help

Start Profiling Run this code:

scorrelator_windowFnc (1 call, 7.782 sec)

Generated 23-Jan-2013 06:35:25 using cpu time.

M-function in file C:\Users\Akber\Dropbox\ch5tests\OTMACH Spatial\scorrelator_windowFnc.m

Copy to new window for comparing multiple runs

Refresh

Show parent functions Show busy lines Show child functions

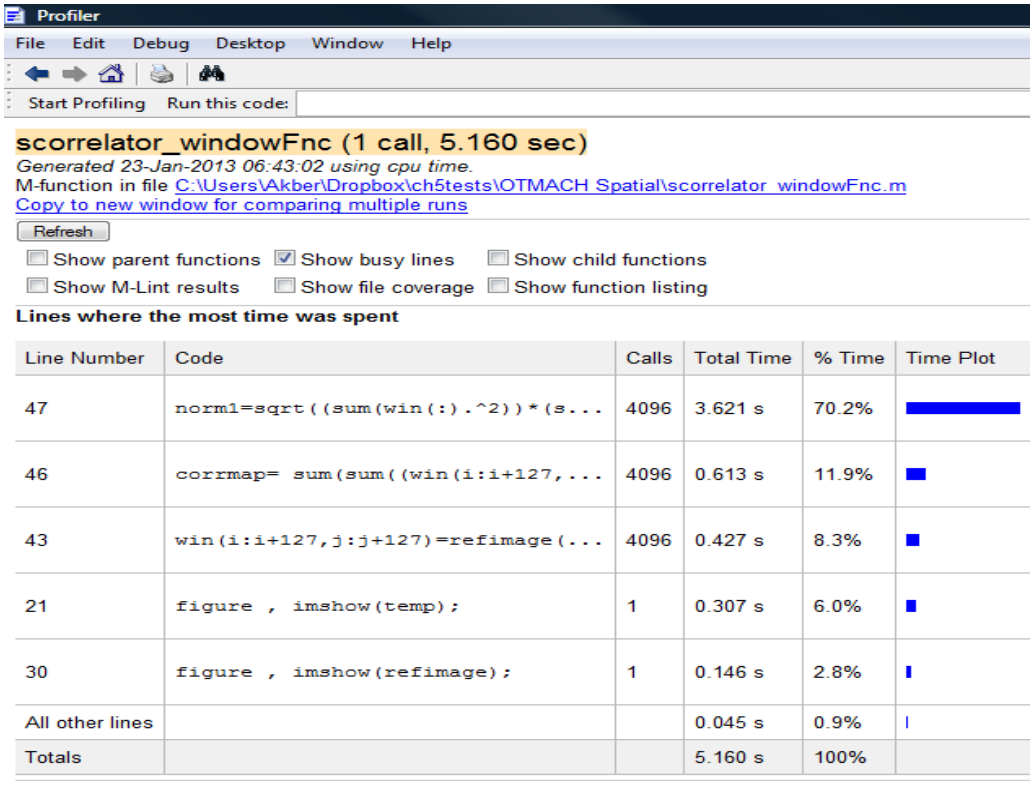
Show M-Lint results Show file coverage Show function listing

Lines where the most time was spent

Line Number	Code	Calls	Total Time	% Time	Time Plot
47	norm1=sqrt((sum(win(:).^2))*(s...	16129	3.310 s	42.5%	
46	corrmap= sum(sum((win(i:i+127,...	16129	2.241 s	28.8%	
43	win(i:i+127,j:j+127)=refimage(...	16129	1.089 s	14.0%	
21	figure , imshow(temp);	1	0.760 s	9.8%	
30	figure , imshow(refimage);	1	0.168 s	2.2%	
All other lines			0.214 s	2.8%	
Totals			7.782 s	100%	

Figure 5.5 Timing Results for the SPOT-MACH filter without application of low-pass enhancement

In Figure 5.5 the timing analysis for the SPOT-MACH without enhancement is given which shows that the correlation process takes around 7.782 seconds to complete the correlation phase which is too slow to be acceptable. With the application of low-pass filtering technique the following speed-up was achieved.



Line Number	Code	Calls	Total Time	% Time	Time Plot
47	norm1=sqrt((sum(win(:).^2))*(s...	4096	3.621 s	70.2%	
46	corrmap= sum(sum((win(i:i+127,...	4096	0.613 s	11.9%	
43	win(i:i+127,j:j+127)=refimage(...	4096	0.427 s	8.3%	
21	figure , imshow(temp);	1	0.307 s	6.0%	
30	figure , imshow(refimage);	1	0.146 s	2.8%	
All other lines			0.045 s	0.9%	
Totals			5.160 s	100%	

Figure 5.6 Timing Results for the SPOT-MACH filter with application of low-pass enhancement

From Figure 5.6 it is seen that the with the application of enhancement technique there was an improvement in performance of the SPOT-MACH filter where the effective time was reduced from 7.782 seconds to 5.160 seconds. Although in the spatial correlator phase due to the reduction in the number of computations there was some timing improvement in the multiply accumulate phase but no viable reduction was noticed.

Although the low-pass filtering based enhancement techniques provides some gain in performance of the SPOT-MACH filter a major problem in this approach is the random nature of regions that are identified for further processing. The low-pass filter will return regions that are within the threshold defined many of which are unwanted. To improve upon this, an alternative approach was developed which not only effectively reduced processing time but also provided an estimate of the location of the object by employing entropy based filtering. This approach is discussed in more detail in the next section.

5.2.2 Speed Enhancement Using an Entropy Based Filter

The use of entropy for background segmentation has been used quite extensively in the past [81]. It is established that the smaller the local entropy, the bigger the threshold of change in that region as described by equation 5.1 and 5.2. Here $f(x,y)$ is the greyscale intensity a pixel (x,y) and $f(x,y)>0$ with $M \times N$ being the image size, than the local entropy E_f is given by [80] , [82]:

$$E_f = -\sum_{i=1}^{M}\sum_{j=1}^{N} p_{ij} \log p_{ij} \tag{5.1}$$

where p_{ij} is the image greyscale distribution which is given by [80]:

$$p_{ij} = \frac{f(x,y)}{\sum_{i=1}^{M}\sum_{j=1}^{N} f(i,j)} \tag{5.2}$$

The local entropy reflects the dispersion of the grey scale intensities of the image, the distribution is uniform in an area where local entropy is big, and the changes in the image greyscale intensities are big where local entropy is small [81].

As has been discussed above, in the case of low pass filtering enhancement, the regions of interest acquired were not very accurate. Based on the concept of entropy, an enhancement technique is proposed which combines the principles of background segmentation and entropy filtering to give regions of interest which are of smaller size and more accurate in terms of location of the target than those provided by the low pass filtering method. Although background subtraction techniques perform well in extracting most of the relevant pixels of moving regions, they are usually sensitive to sudden changes such as variations in illumination. For this reason, the developed enhancement technique is not a traditional background segmentation technique but is rather an estimation technique which uses difference of entropies calculated over a moving window to predict the location of a target and tackles illumination changes using the SPOT-MACH filter in a two staged approach. The developed approach is summarised in Figure 5.7 shown below.

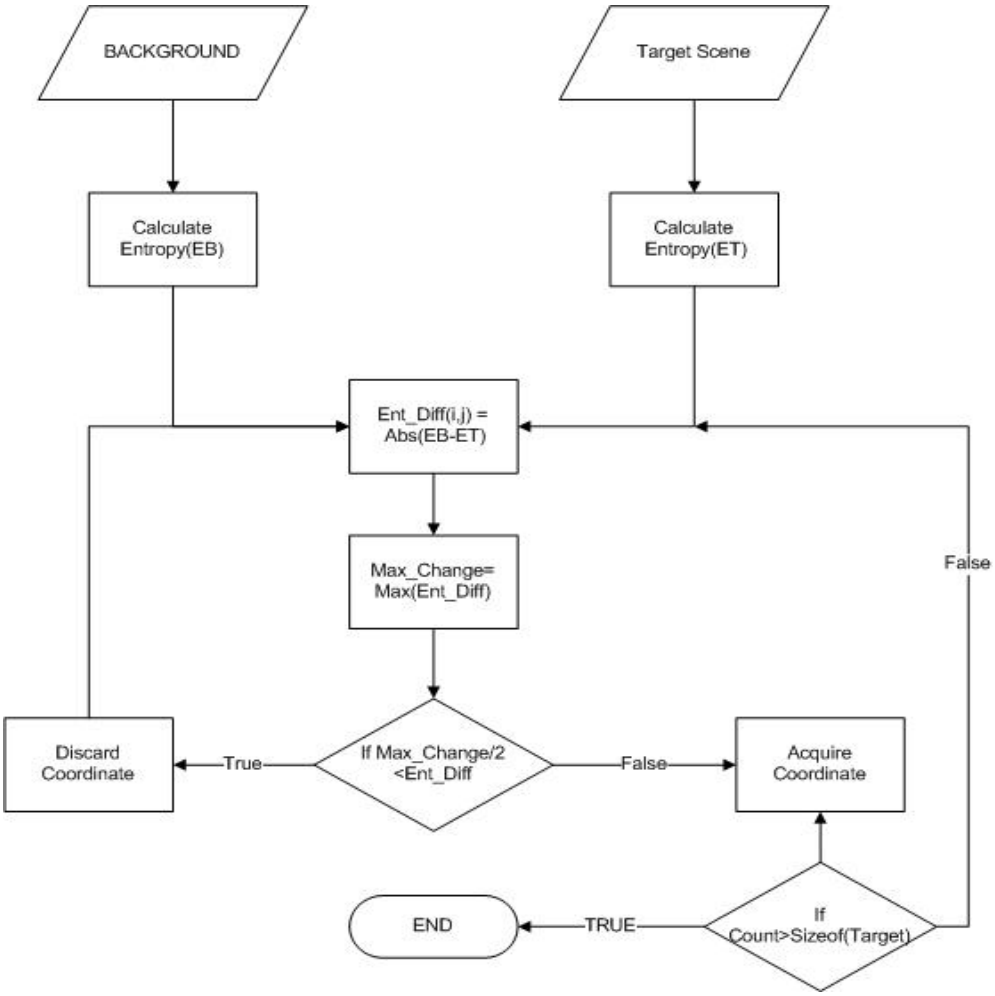
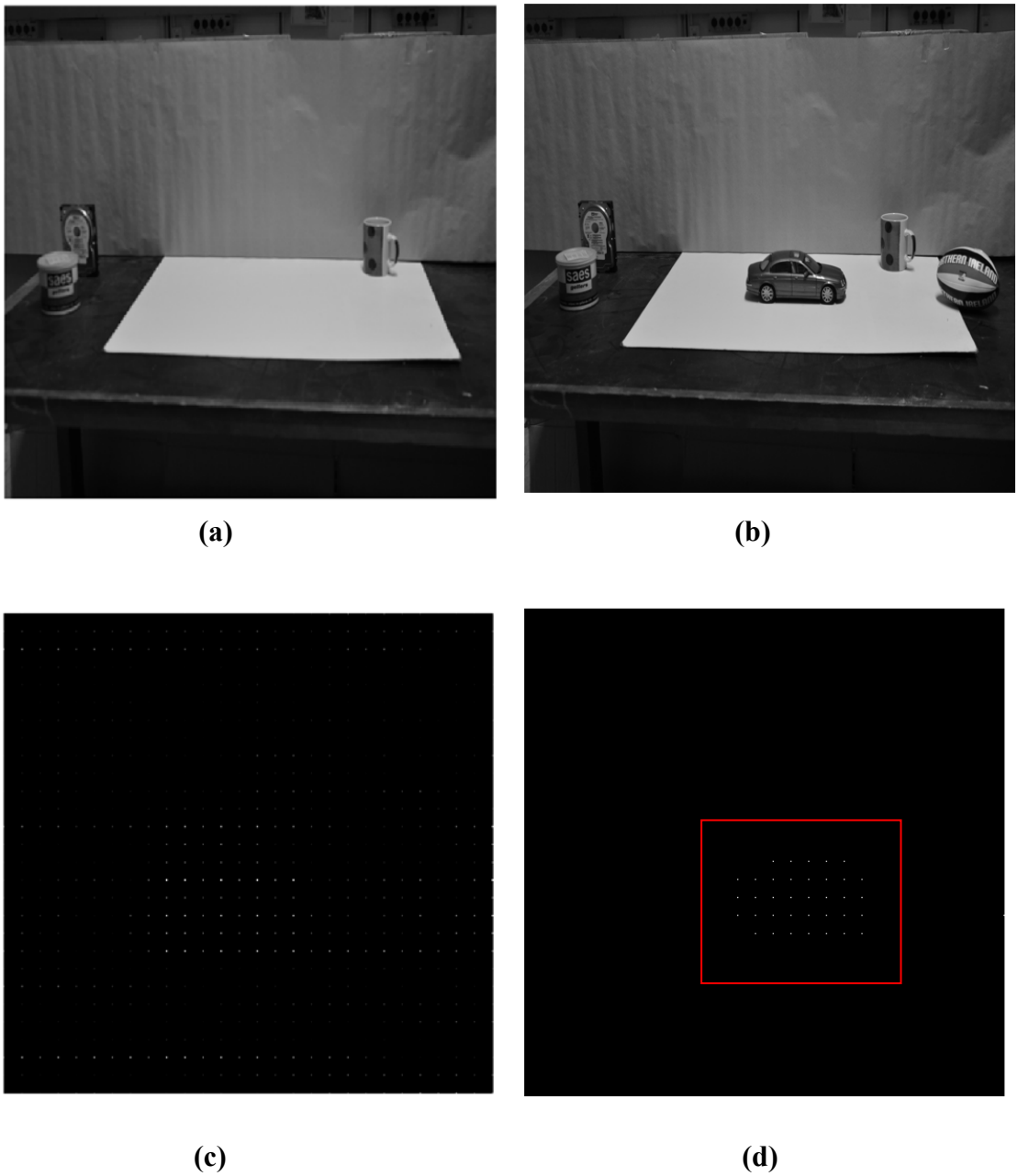


Figure 5.7 Entropy filter based enhancement technique for the SPOT-MACH filter

From Figure 5.7 the working of the entropy based enhancement technique is given, which shows that the entropy of the background and the target object are calculated and than their difference is found. The maximum of the entropy difference divided by 2 is considered the threshold any entropy value below this is rejected and the values above this are assigned a pixel intensity value 255 so that the region could be identified.

The acquired coordinates of the region of interest are passed on to the spatial correlator. The spatial correlator is modified in terms of the movement of the windowing kernel. Instead of starting the kernel from zero the spatial correlator positions the kernel on the smallest coordinate and uses shifts of five pixels up , down , right and left to minimise correlation loss.



(a) **(b)**

(c) **(d)**

Figure 5.8 Background Image (a), Target Image (b), Entropy Difference (c) and Region of Interest after threshold function (d)

In Figure 5.8 the application of the entropy based enhancement technique is shown, it can be seen that at the end the region of interest contains only 34 points and the location of the target lies within the region as well. A detailed timing analysis of the entropy based technique is summarised in Figure 5.9.

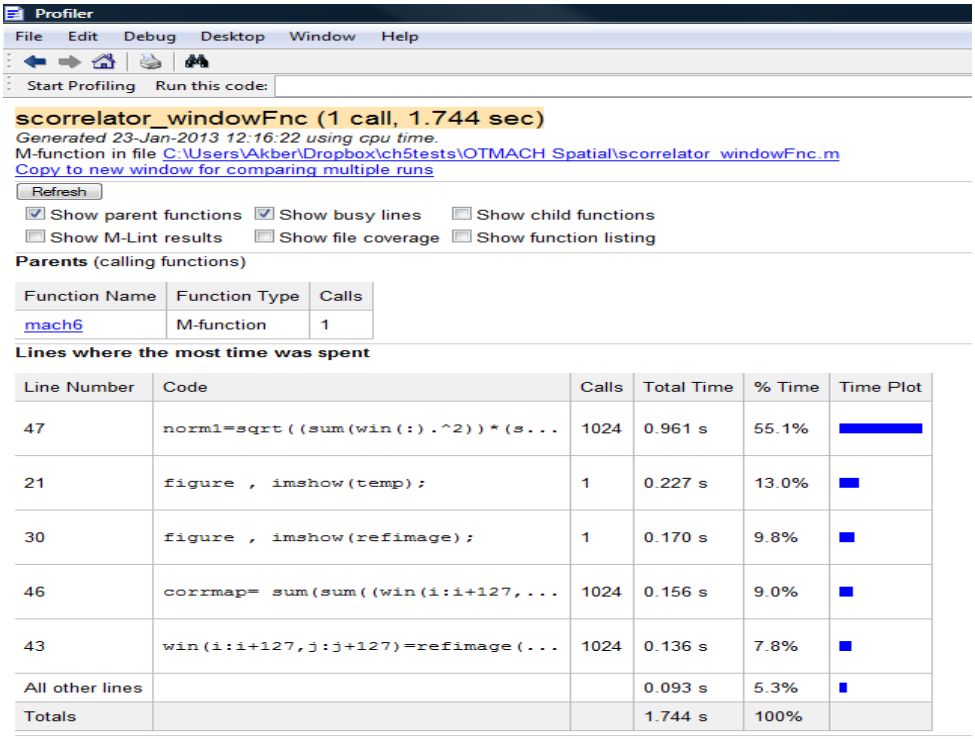


Figure 5.9 Timing analysis for entropy based enhancement technique

In Figure 5.9 it can be seen that after the application of entropy based enhancement technique the timing of the SPOT-MACH filter is comparable to the frequency domain OT-MACH filter. The correlation process which is the most time consuming process in the SPOT-MACH is taking 0.961 seconds whereas the same process in the frequency domain was taking 0.871 seconds as discussed earlier so hence there is a difference of 0.09 seconds between the frequency domain and spatial domain techniques which can be considered negligible. Also in SPOT-MACH the enhancement technique has also effectively predicted the location of the target.

The enhancement technique discussed in this case is a software based enhancement technique further enhancements can be achieved in the spatial domain by employing hardware based techniques which are discussed further on in this chapter.

5.3 Hardware Based Optimization Techniques

In the previous section two optimisation techniques were discussed, both of which were software based approaches. In this section, hardware based approaches are developed using a GPU and a FPGA to optimise the timing performance of the SPOT-MACH filter.

5.3.1 GPU based Optimization Technique

GPU computing is the use of a GPU together with a Central Processing Unit (CPU) to accelerate general-purpose scientific and engineering applications. Pioneered five years ago by NVIDIA, GPU computing has quickly become an industry standard, enjoyed by millions of users worldwide and adopted by virtually all computing vendors [50] , [83].

GPU computing offers unprecedented application performance by offloading compute-intensive portions of the application to the GPU, while the remainder of the code still runs on the CPU. The use of CPU combined with GPU is a powerful combination because CPUs consist of a few cores optimized for serial processing, while GPUs consist of thousands of smaller, more efficient cores designed for parallel performance. Serial portions of the code run on the CPU while parallel portions run on the GPU.

In order to develop the GPU based optimization technique a spatial correlation routine was written in using C and the following libraries were added to access the GPU functionality ‘QtCore/QCoreApplication’, ‘opencv2/opencv.hpp’ and ‘opencv2/gpu/gpu.hpp’. The correlator routine developed was run on the CPU as well to assess the performance gain of the GPU as compared to the CPU.

The choice of GPU for this implementation was the ‘Quadro 4000’ professional graphics solution, built on the NVIDIA Fermi architecture, used in a broad range of design, animation and video applications. Modern applications harness the NVIDIA CUDA parallel processing architecture of the ‘Quadro’ GPU to deliver performance gains up to eight times faster than previous generations when running computationally intensive applications such as ray tracing, video processing and computational fluid dynamics. The choice of CPU for the comparison was the Intel(R) Core(TM)2 Quad

CPU Q6600 running at 2.40 GHz. A comparison chart was created between the two processors when the experiment was carried out on a 4096 × 4096 pixel image using multiple window sizes of 64 × 64 , 128 × 128 , 512 × 512 and 1024 ×1024. The GPU speed up factor comparison is shown in Figure 5.10 below..

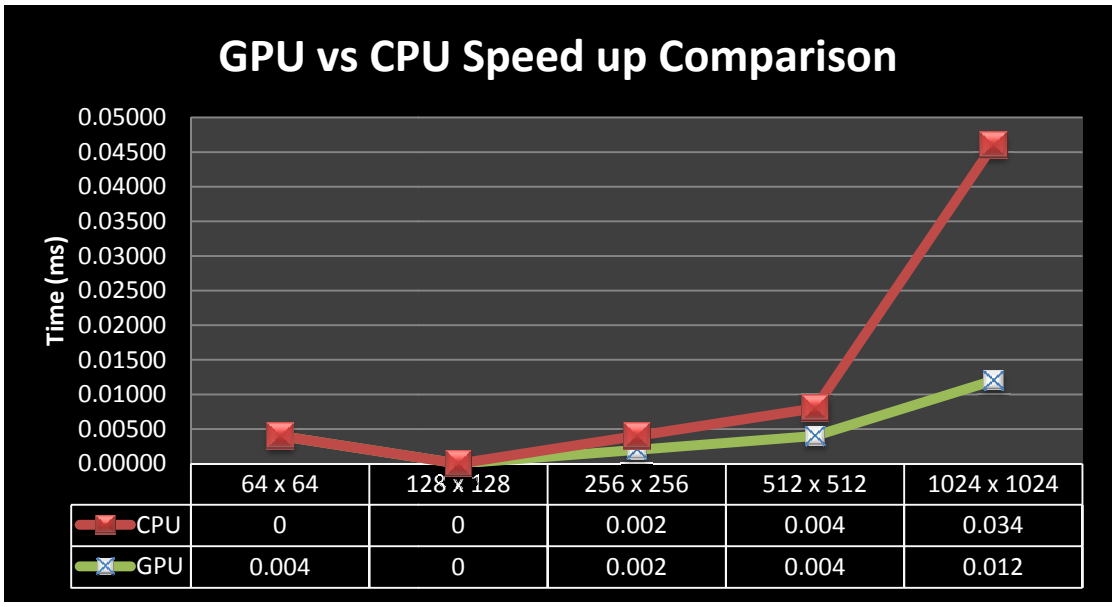


Figure 5.10 GPU versus. CPU speed up comparison using the Spatial Correlator

From Figure 5.10 it can be seen that the GPU performance is dependent on the size of the data window. In this case up till 128x128 the GPU shows no speed-up but as soon as it crosses 256×256 window size the GPU performance starts to increase and eventually speed up was achieved for bigger windows.

In the case of GPU the speed up achieved is due to the parallel processing cores that can be utilized but the problem of using multiple cores for small datasets is that it creates bottle necks. The bottle neck is created at the memory read\write operation which is not dependent on the GPU but rather the memory bus of the system. A comparison of GPU with the memory read write cycle delays is given by Figure 5.11 below.

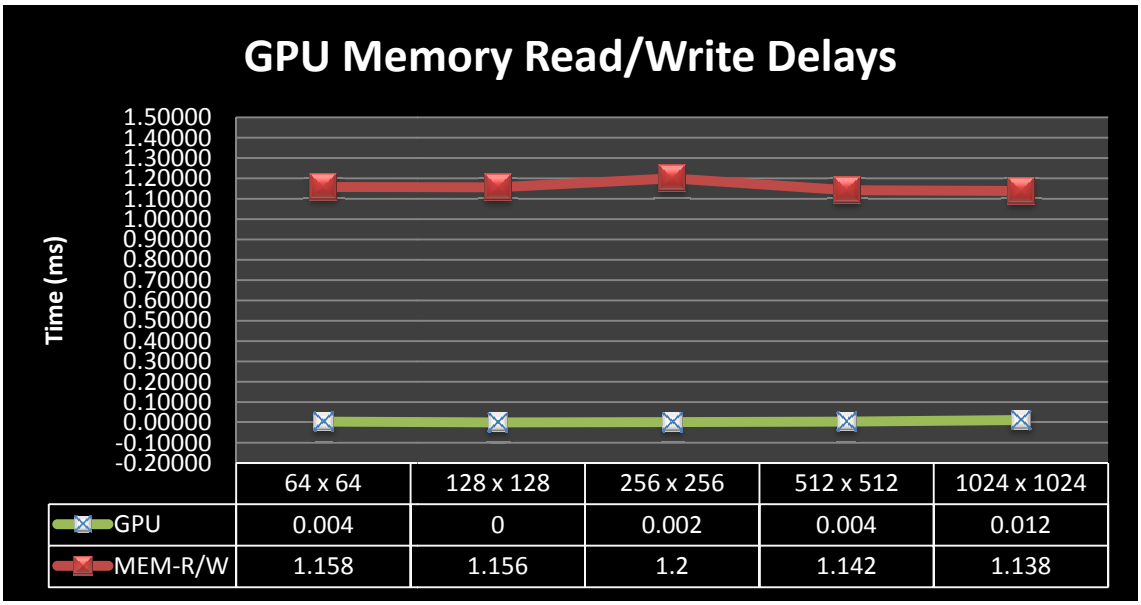


Figure 5.11 GPU versus Memory read\write delays

From Figure 5.11 it can be seen that the memory read\write latencies account to a lot delay in the GPU processing. There is a lot of speed up achieved from the GPU but the data latencies account for the memory read\write time loss. In the case of GPU there are more chances of data dependencies for smaller data sets.

5.3.2 FPGA based optimization technique

In order to design a system earlier say 10 – 20 years back the manufacturer had to bear a very high cost due to the fact that rapid testing and re design was constantly required. In the manufacturing industry Application specific integrated circuits (ASIC) circuits are used to form non reconfigurable but fast systems being manufactured on silicon wafers of a specific die.

However the cost of designing ASICs is increasing every year. In addition to the non-recurring engineering (NRE) and mask costs, development costs are increasing due to ASIC design complexity. Issues such as power, signal integrity, clock tree synthesis, and manufacturing defects can add significant risk and time-to-market delays. To overcome the risk of recalls, high NRE costs, and to reduce time-to-market delays, FPGAs offer a viable and competitive option to ASIC development.

FPGA applications have led to higher density devices, intellectual property (IP) integration, and high-speed I/O interconnects technology. All of these elements have allowed FPGAs to play a central role in the defence and security industry.

In digital design an n-bit lookup table can be implemented with a multiplexer whose select lines are the inputs of the LUT and whose inputs are constants. An n-bit LUT can encode any n-input Boolean logic function by modelling such functions as truth tables. This is an efficient way of encoding Boolean logic functions, and 4-bit LUTs are in fact the key component of modern FPGAs.

In order to implement SPOT-MACH in FPGA sequences of steps have to be followed. The design tools used were ISE “PROJECT NAVIGATOR” by “XILINX” and for simulating “MODEL SIM” by “MENTOR”.

The reason behind choosing ISE PROJECT NAVIGATOR is due to the fact it has an extensive choice of tools that range from synthesis to place and route. Also it has options for manual place and route. The project navigator has specific libraries for the SPARTAN XC2S 200E it also has “IMPACT” which enables JTAG flashing of the boot rom. “IMPACT” runs a boundary scan and checks for connected devices and shows a virtual view of the device.

The simulation tool “MODEL SIM” is one of the most widely used tool in the industry for its relatively flexible system requirements as opposed to CADENCE tools.

A block diagram for the design of the SPOT-MACH for FPGA implementation is given by Figure 5.12 below.

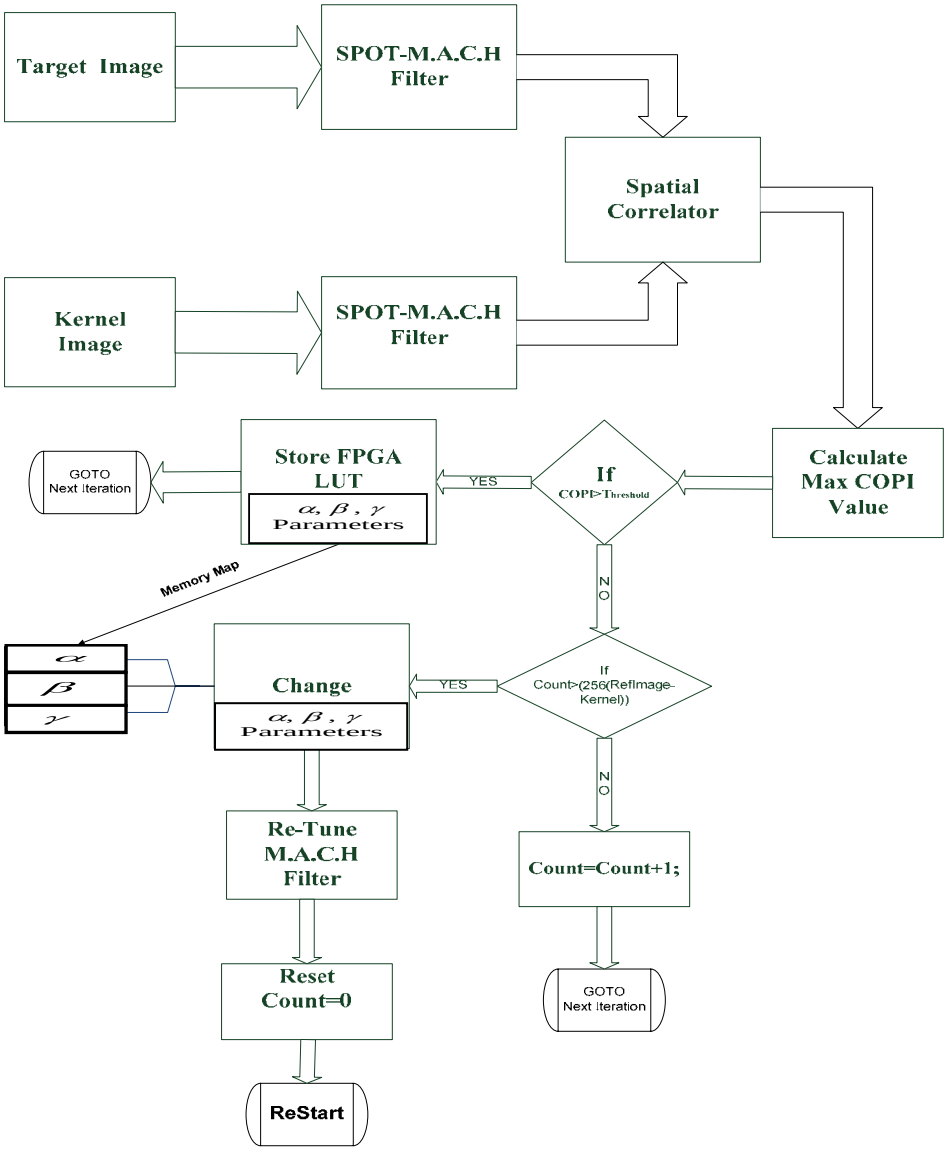


Figure 5.12 FPGA based enhancement technique for SPOT-MACH

In Figure 5.12 the hardware implementation of a SPOT-MACH is given. It is shown that when the kernel and the input images are acquired they are first pre-processed through an SPOT-MACH filter and then passed on the spatial correlator. The spatial correlator comprises of multiply accumulate unit which normalizes the kernel and performs cross correlation on the two images.

The optimum SPOT-MACH parameters are stored in the lookup table, the system keeps on finding new optimized values for the non negative parameters to improve target recognition.

The system has been tested by running tests on a SPARTAN XC2S200E FPGA. The performance evaluation criteria was defined by keeping into account the processing time and resources required during post place and route simulation on a FPGA.

Now when the enhancement algorithm was designed on a SPARTAN-II XC2S200E FPGA the timing summary generated by "ISE PROJECT NAVIGATOR" is given in Table 5.1 below.

Min.period	Min. input arrival time before clock	Max. output time after clock:	Maximum combinational path delay
13.219ns	8.840ns	14.126ns	11.87ns

Table 5.1 SPARTAN XC2S200E timing summary

The overall design frequency in this case is 75.69 MHz and by comparing the clock speeds it can be seen that the FPGA implementation is considerably faster.

Also 'PROJECT NAVIGATOR" gives the option of design optimization as well. The design can be optimized according to speed or area used on the FPGA. This can be shown in Figure 5.13 given below.

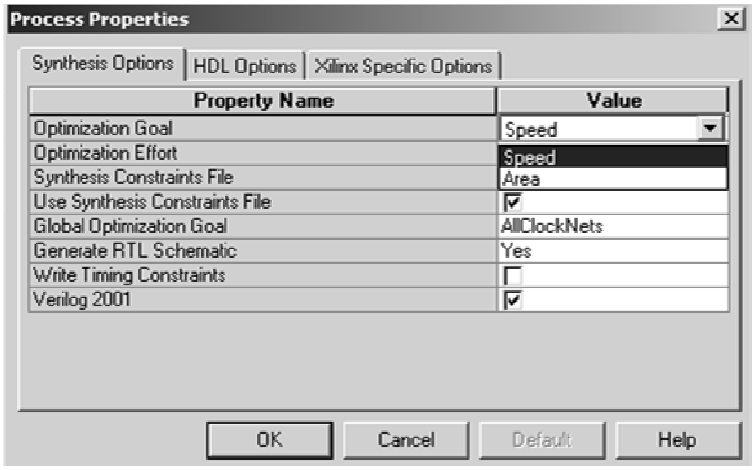


Figure 5.13 SPOT-MACH FPGA design optimization

The chip layout for the FPGA can be seen in Figure 5.12 given below.

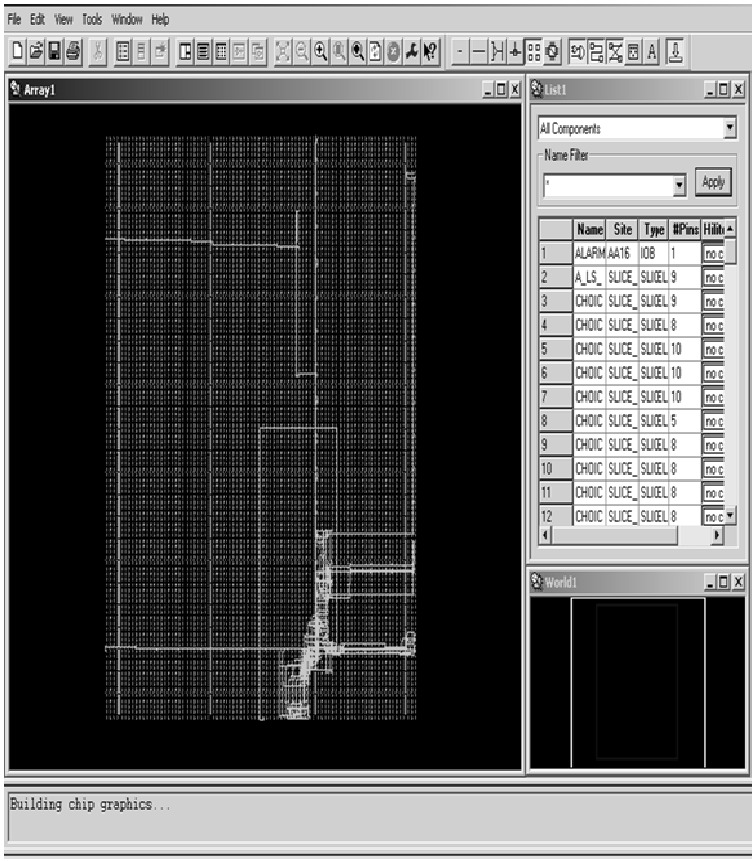


Figure 5.12 SPOT-MACH FPGA chip layout

It can be seen in Figure 5.12 the multiplier and adders have taken a small area on the chip, smaller the area required quicker the data bus speed. With the rapid advances in the digital domain it's only a matter of time when it overcomes the optical domain.

Considering the speed of the FPGA it has been proposed as an ideal choice for implementing the SPOT-Filter enhancement.

The most ideal choice would be to use a VIRTEX-5 FPGA for hardware implementation which has a clock speed of 550 MHz and can compute up to 528 Giga Multiply Accumulate functions. Also it has a DSP48E slice giving it the additional edge of having a DSP processor which has dedicated 25x18 multipliers.

5.4 Conclusion

In this chapter a detailed timing analysis was performed and two software and hardware based techniques were proposed. Low-pass filters were used to define regions of interests as was an entropy based filter, the later proving to more accurately locate potential target regions in the image. The computationally intensive nature of the space domain implementation may be further tackled by the use of an efficient digital hardware implementation using FPGAs. A performance evaluation has been performed on the SPOT-MACH filter by comparing the computational results from an INTEL CORE 2 DUO processor and a SPARTAN XC3S200E board. In this chapter it was established that the SPOT-MACH filter, along with the enhancement techniques described, is the most viable approach to overcome the problem of illumination changes.

.

Chapter 6

Applications of SPOT-MACH and Comparison with Feature Based matching techniques

6.1 Introduction

In the previous chapters the design and implementation of the SPOT-MACH filter along with its speed enhancement techniques have been discussed. It has also been established in the previous chapters that the SPOT-MACH filter can not only be designed to be invariant to change in orientation of the target object but also to be spatially variant, i.e. the filter function can be made dependent on local clutter conditions within the image. Sequential location of the kernel in all regions of the image does, however, require excessive computational resources. In order to overcome the computational overhead of the SPOT-MACH filter some speed enhancement techniques have also been proposed in the previous chapter.

A number of non-parametric local regression techniques have also been described that offer a simplified approach to pattern recognition problems which employ linear filtering using low pass filters designed using moving window local approximations. In most of these cases the algorithms search for a region of interest near the point of estimation for various prevailing conditions which fit the required criteria. These estimates are calculated for a defined window size which is determined as being the largest area within which the estimators do not widely vary from the criteria [80] , [78].

The only drawback in this approach is that the window size is directly proportional to the required computational resources and would adversely affect the performance of the system if the moving window size is not proportionate to the resources. One of the optimization techniques which have been discussed in Chapter 5 employs low-pass

filtering to highlight the potential region of interests in the image and then restricts the movement of the kernel to these regions to allow target identification. Also another proposed highly effective technique employs a proposed entropy matrix filter to identify potential regions of interest to reduce the number of shifts for the kernel. A GPU based enhancement technique has also been discussed elaborating the gain in performance of the SPOT-MACH if the computations were carried through a GPU but this technique was found to be ineffective for the desired size of the kernel being used but for larger kernels has the potential to offer performance gain. Also a FPGA based implementation technique has also been discussed highlighting the potential advantages in speed enhancement by utilizing the built in multipliers and look up tables (LUT) within the FPGA architecture [84].

In this chapter the target detection and identification capability of the proposed two-stage process using SPOT-MACH has been compared in highly cluttered backgrounds using both visible and thermal imagery and associated training data sets. A performance matrix comprised of peak-to-correlation energy (PCE) and peak-to-side lobe ratio (PSR) measurements of the correlation output has been calculated to allow the definition of a recognition criterion. Also in this chapter the subsequent capabilities of the SPOT-MACH correlation filter have been compared with Scale Invariant Feature Transform (SIFT) which is a popular feature based detection technique. Based on these comparisons the capabilities of the proposed SPOT-MACH filter have been assessed in terms of ability to detect target, suppression of noise and finding the location of the target in various illumination conditions.

6.2 Applications of the SPOT-MACH filter using FLIR imagery

An infrared sensor can detect infrared radiation and by converting the temperature differences between the object and the surroundings creates a temperature scale. The temperature scale is converted into a greyscale map and an image is obtained. The images obtained from such a sensor enables the detection of an object through smoke in a burning building, adverse weather conditions or in the absence of any reflected light. The most common type of sensor used for acquiring such images is known as Forward

Looking Infrared Sensor (FLIR) and images obtained from it are referred as FLIR images [35] , [85] and [86].

The detection and recognition of targets in FLIR images has always been a challenging problem due to the varying heat signature of the object and background clutter. The FLIR imagery used in this chapter has been acquired from a moving platform and contains multiple objects at different orientations and background variations. The movement of the sensor and the object induce coupled motions into the FLIR images which make the detection and tracking of the target object extremely complicated due to these varying criteria, Figure 6.1 shows an example view of a vehicle with a hot engine but cool passenger compartment. Considering all these factors, in particular the variable nature of the source emission patterns in FLIR images, it was decided to assess the SPOT-MACH filter with this imagery.

Figure 6.1 Example of a FLIR image acquired in a desert terrain

As established in Chapters 4 and 5 the SPOT-MACH filter is a very powerful correlation filter; it offers reliable detection of the correlation peak, distortion tolerance, ability to suppress clutter noise and is tolerant to illumination changes.

The FLIR imagery used in this thesis has been provided by the courtesy of the Kuwait Ministry of Defense (MOD) and covers a harsh desert terrain for testing purposes. A 3D model of a Nissan Patrol vehicle (NP) oriented from 0 to 360 degrees with a 10 degree

shift between every sample has been used to provide training images as shown in Figure 6.2.

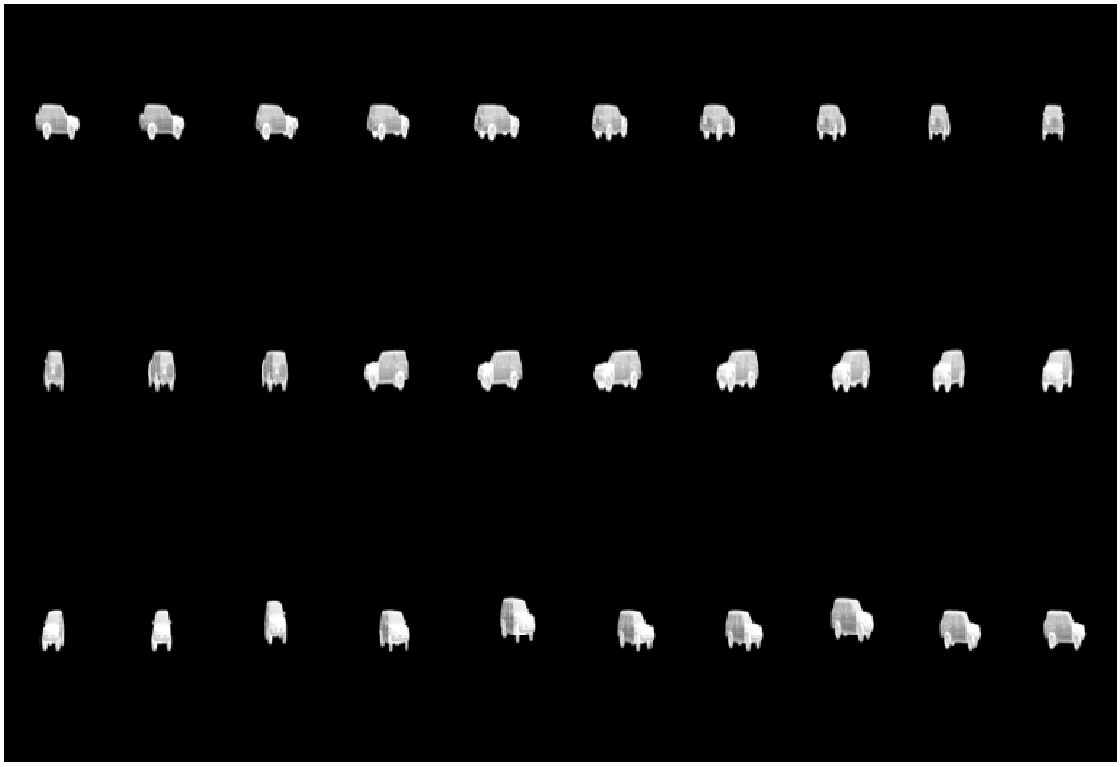


Figure 6.2 A 3D Model of Nissan Patrol (NP) for Training Image Dataset

The 3D model of the NP was created using actual design ratios to increase the chances of target recognition. The training images have been tested with real life FLIR imagery of a Nissan Patrol vehicle entering a secured perimeter in harsh weather conditions to evaluate the detection capabilities of the filter. It was shown that an appropriately SPOT-MACH filter can be made invariant to some range of in-plane and out-of-plane rotations and can be tuned to produce defined detection criteria.

In this section, the SPOT-MACH has been implemented and the performance evaluated in different scenarios relating to perimeter breaches for security applications. The FLIR imagery used covers a harsh desert terrain which is ideal for testing purposes as it provides varying heat signatures of both the target vehicle and the clutter background.

From Chapter 4 it has been established that SPOT-MACH filter allows some degree of distortion invariance and can yield sharp detection peaks in the presence of limited degrees of in-plane and out-of-plane rotation.

The initial test conducted for the work shown in this chapter using the Nissan Patrol vehicle 3D model oriented at zero degrees correlated with the same image with no background noise to evaluate the ideal correlation filter response, as shown in Figure 6.3.

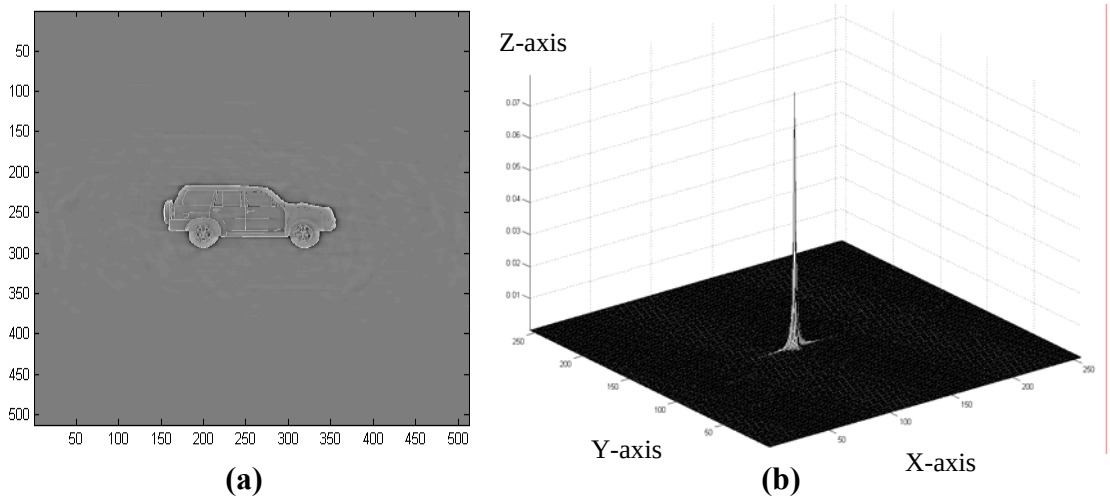


Figure 6.3 NP oriented at 0 degrees (a) and output correlation plane (b)

The correlation plane shown in Figure 6.3(b) has a sharp peak at the center which shows that the correlation filter successfully matched the reference object with the target object, in this case the NP image. The correlation plane in Figure 6.3(b) has the following performance ratios as shown in Table 6.1:

COPI	PSR	PCE	α	β	γ
0.0795	102.12	0.1603	0.0001	0.2	0.001

TABLE 6.1 Performance ratios for NP oriented at 0 degree

The performance ratios shown in table 6.1 indicate the ideal COPI and PSR values for the successful recognition of the target object. Now the designed filter must be tested with different orientations and with cluttered backgrounds to evaluate the performance.
In order to test the discrimination abilities of the correlation filter in the presence of background clutter, tests were conducted with NP oriented at zero degrees. The background clutter used had a desert terrain with a vehicle moving on a dirt road with varying heat signatures. The correlation plane with the reference and target images can be seen in Figure 4 given below.

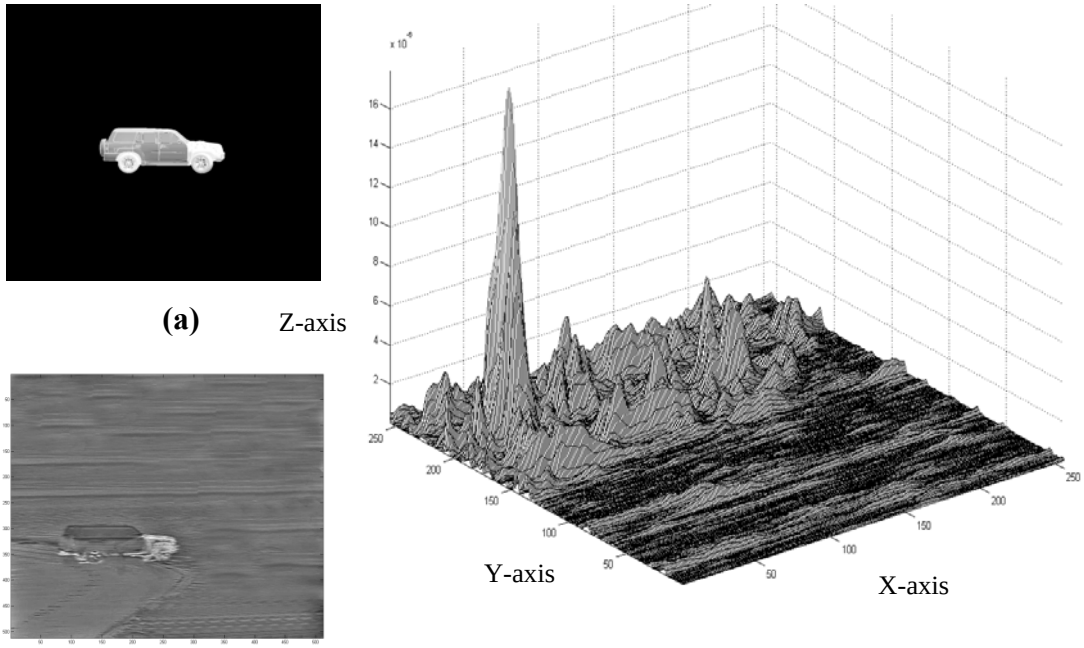


Figure 6.4 The NP rotated at 0 degrees (a), FLIR image with background clutter (b) and correlation output plane (c)

Figure 6.4 contains a reference object which is a NP oriented at zero degree, a target object which is a FLIR image with a moving NP and the output correlation plane. The correlation plane illustrated in Figure 6.4 is built up pixel by pixel, yielding the localized correlation response. It can be seen that although with varying heat signatures, which renders the reference object almost indistinguishable from the background, there is a visible correlation peak which signifies that the object is detected. This can be better

quantified with the performance ratios given in Table 6.2 for the correlation output plane shown in Figure 6.4(c).

COPI	PSR	PCE	α	β	γ
$1.78*10^{-5}$	9.43	0.0014	0.0001	0.2	0.001

TABLE 6.2 Performance ratios for NP oriented at 0 degree with background clutter

From Table 6.2 it can be seen that the value of PSR has dropped due to the background clutter as there are multiple peaks and there is a drop in the COPI value as well. But even with these prevailing conditions the filter was able to distinguish the object readily.

In the above experiments the ability of the filter to discriminate an object from cluttered background was tested. Although the filter managed to detect an object in the presence of clutter there was a significant drop in the performance ratios. This was due to the fact that the previous experiment was conducted using a single training image and the filter was trained at single degree of orientation. So in order to better evaluate the abilities of the filter a new set of tests were conducted with a different set of FLIR images containing a different pattern of background noise.

The target object was oriented at different angles and location with the training images multiplexed at multiple angles. Figure 6.5 shows an example which contains a background image with the target centered within extreme clutter together with a set of nine multiplexed training images oriented between 150 and 230 degrees with ten degree increments. The parameters used for the OT-MACH filter are kept the same as in the previous experiments conducted.

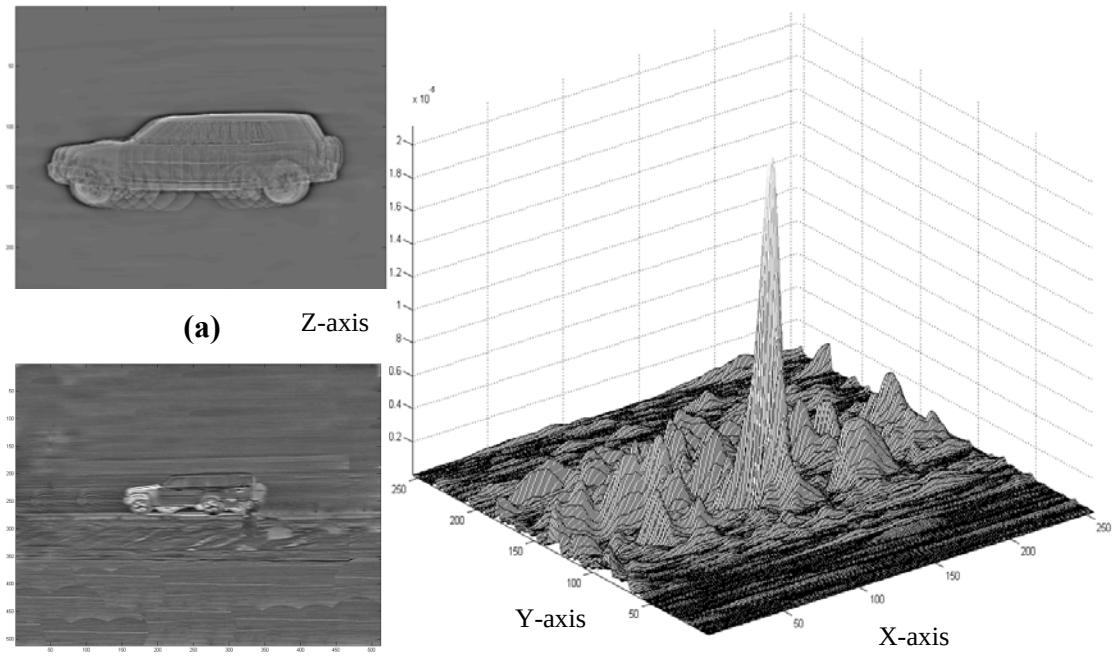


Figure 6.5 NP rotated between 150-230 degrees (a), FLIR image with background clutter (b) and correlation plane (c)

In Figure 6.5, the filter's ability to successfully discriminate a target oriented at 180 degrees within background clutter is shown to improve when trained with multiple orientations of the NP data set. This can be quantified from Table 6.3 given below.

COPI	PSR	PCE	α	β	γ
2.78*10^-5	9.29	0.0017	0.0001	0.2	0.001

TABLE 6.3 Performance ratios for NP oriented at 150-230 degree with target oriented at 180 degree

It is shown in Table 6.3 that by increasing the number of training images allows the filter to perform better by yielding sharper peaks and higher PSR and PCE values.

Another interesting scenario is given in Figure 6.6 where a NP is moving in the range of the FLIR sensor and the filter has to discriminate against background clutter plus camera information feeds, resulting in further noise. In this case the NP is oriented at an angle of 30 degrees in the target image with the reference image containing multiple orientations between 150 and 250 degree.

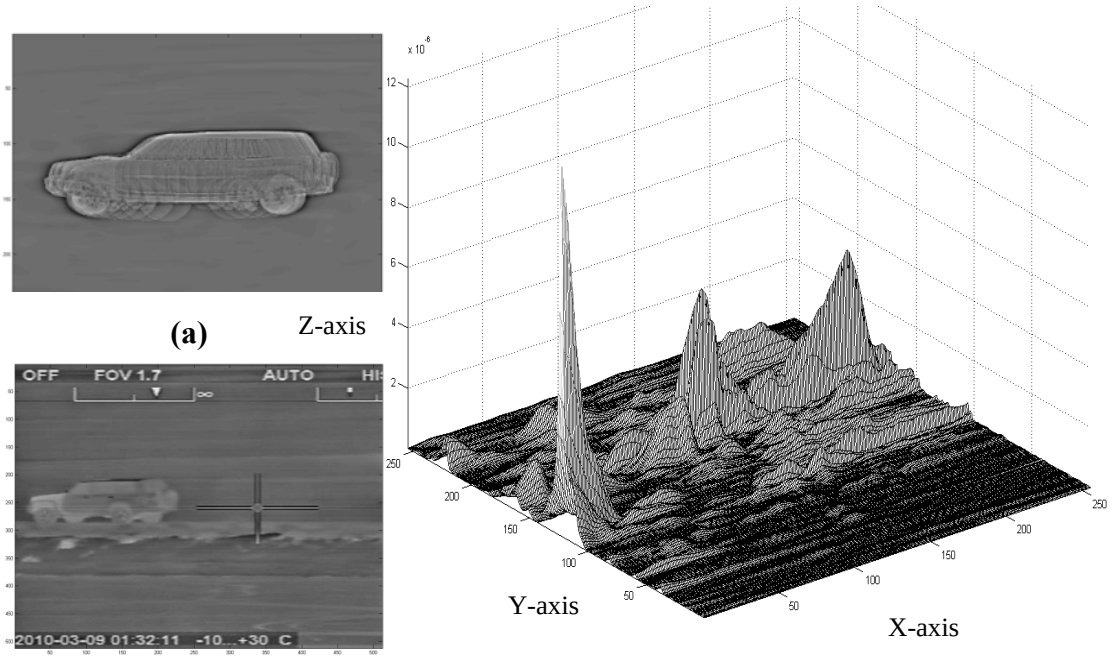


Figure 6.6 NP rotated at 150-250 degrees (a), FLIR image with background clutter (b) and Correlation Plane (c)

It can be seen in figure 6.6 that there is a sharp peak at the location of the target object but there are also visible side lobes. The performance ratios for the correlation plane in Figure 6.6(c) are given in Table 6.4 below:

COPI	PSR	PCE	α	β	γ
9.28*10^-5	10.29	0.002	0.0001	0.2	0.001

TABLE 6.4 Performance ratios for NP oriented between 150-250 degree with target oriented at 180 degree with additional camera noise

It can be seen from Table 6.4 the performance ratios of the correlation plane would enable the detection of the target object. In order to improve the performance of the filter, tests were conducted with additional orientations covering the whole orientation plane between 0 to 360 degrees. The training images used are shown in Figure 6.7 given below:

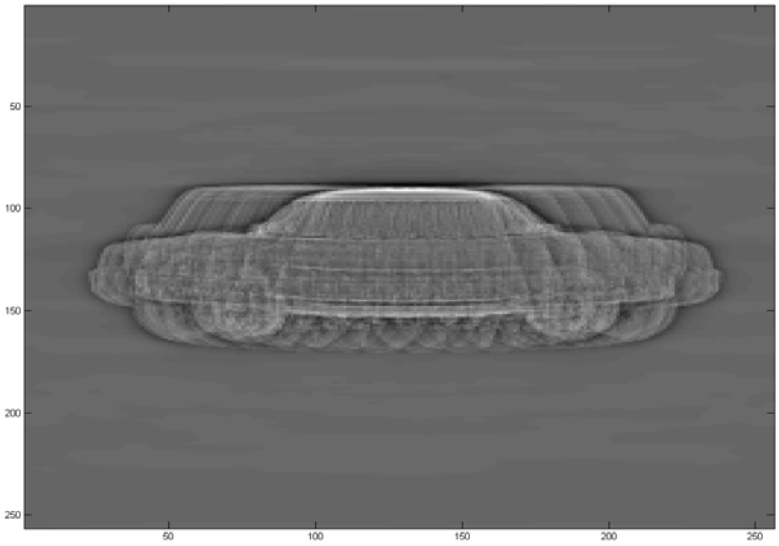


Figure 6.7 NP oriented at 0-360 degrees

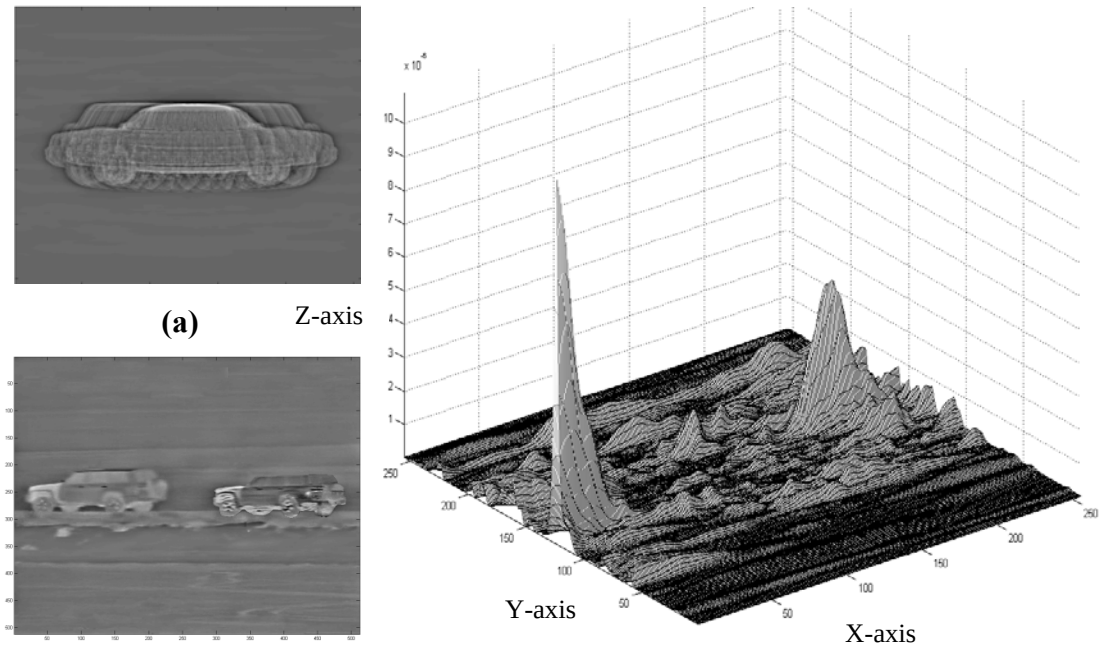
The new sets of results achieved with the added orientations are given in Table 6.5 below.

COPI	PSR	PCE	α	β	γ
4.8*10^-5	10.1	0.0032	0.0001	0.2	0.001

TABLE 6.5 Performance ratios for NP oriented between 0-360 degree with target oriented at 180 degree with camera noise

Comparing Table 6.4 and Table 6.5 it can be clearly seen that there is an increase in the COPI and PSR of the correlation plane with the increase in orientations which enhances the performance of the filter.

The performance of the correlation filter has also been tested by introducing another class of object near the target object to evaluate the discrimination abilities of the filter in scenarios where one needs to classify between permitted and non-permitted classes in a secure perimeter. The results can be seen in Figure 6.8 given below.



(a) **(b)** **(c)**

Figure 6.8 NP rotated at 0-360 degrees (a), FLIR image with multiple targets (b) and Correlation Plane (c)

It can be seen from Figure 6.8 that even in the presence of two different classes of targets where one is a false target the filter has the ability to discriminate correctly and give a sharp a peak in the location of the actual target.

The performance ratios for the correlation plane in figure 6.8 are given in Table 6.6 given below.

COPI	PSR	PCE	α	β	γ
3.8*10^-5	8.5	0.0016	0.0001	0.2	0.001

TABLE 6.6 Performance ratios for NP oriented at 0-360 degree with target oriented at 180 degree with false object

One of the main advantages of the spatial domain implementation is the ability to apply normalisation. It enables adaptation of the filter dependant on background heat signature variances and also enables the normalisation of the filter energy levels. The kernel can be normalized to remove a non-uniform brightness distribution if this occurs in different regions of the image. Another important performance improvement technique that can

be applied on the SPOT-MACH while experimenting with FLIR imagery is the application non-linearity.

6.2.1 Application of a non-Linearity function

As discussed earlier that one of the main advantages of the spatial domain implementation is the ability to apply normalisation. It enables adaptation of the filter dependant on background heat signature variances and also enables the normalisation of the filter energy levels. The kernel can be normalized to remove a non-uniform brightness distribution if this occurs in different regions of the image.

Another important advantage of the spatial domain implementation is application of a non-linearity which releases the energy locked in sharp edges and minimized the chances of false detections [85]. Although the application of non-linearity is dependent on the nature of the images and normally if there are sharp edges present in an image the energy would tend to get locked up at corners.

In order to apply non-linearity the filter transfer function is multiplied by a sigmoid function which is a bounded differentiable real function that is defined for all real input values and that has a positive derivative dependent on two parameters a and c [87].

$$f(x,a,c) = \frac{1}{1+e^{-a(x-c)}} \tag{6.1}$$

Depending on the sign of the parameter a, the sigmoidal function is oriented to the right or to the left.

A wide variety of sigmoid functions have been used as the activation function of artificial neurons, including the logistic and hyperbolic tangent functions. Sigmoid curves are also common in statistics as cumulative distribution functions. An example of a sigmoid curve is shown in Figure 6.9 below.

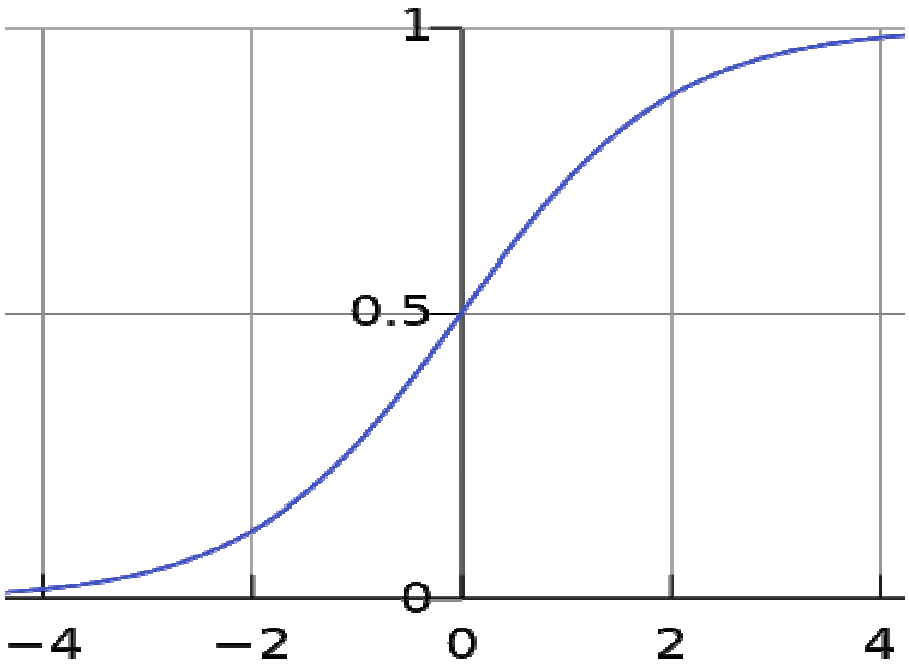


Figure 6.9 Sigmoid Curve for a, c parameters [-2 0]

The application of non-linearity was tested using a brightly illuminated square, the image was processed through the SPOT-MACH transfer function and the resulting image was plotted to observe the energy locked in the edges. This is shown by Figure 6.10 below.

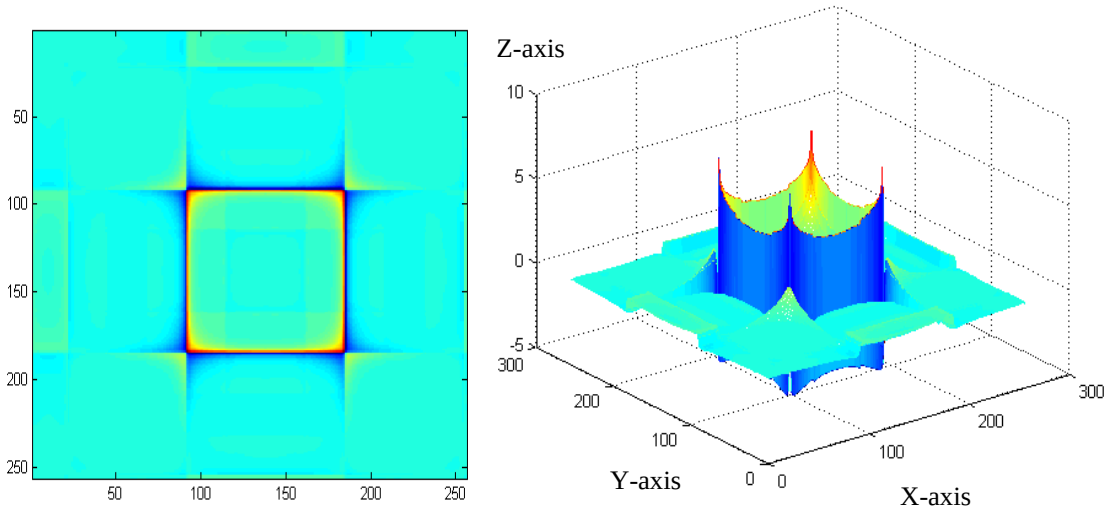


Figure 6.10 SPOT-MACH transfer function for brightly illuminated square

From Figure 6.10(b) it can be seen that the brightly illuminated square as a lot of energy locked at the edges which would could result in false detections. When the sigmoid function was applied to the transfer function the values [0.1 1] for equation 6.1 the following results were observed as shown by Figure 6.11.

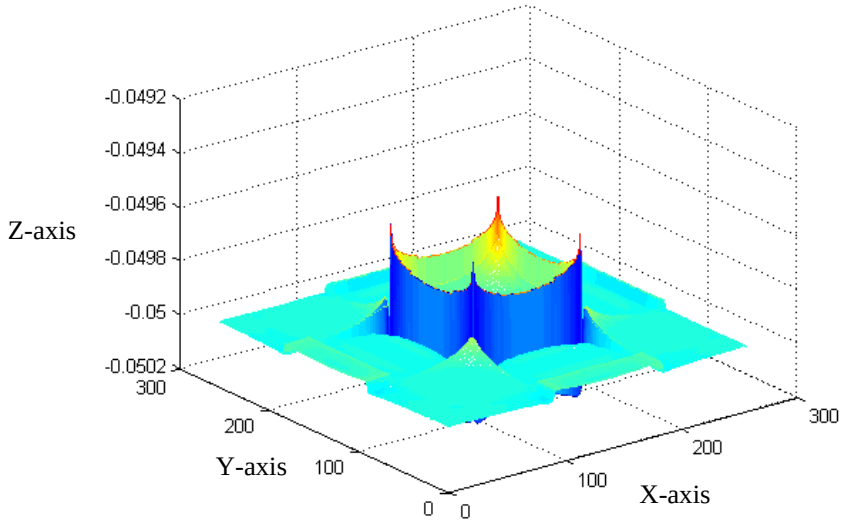


Figure 6.11 SPOT-MACH impulse response after application of non-linearity

It can be observed that the energy at the edges has reduced but a problem arises in that the response of the filter has become negative valued due to the fact that the non-linearity function was reversed and so once the autocorrelation output plane is calculated it becomes reversed as well. This is shown in Figure 6.12 below.

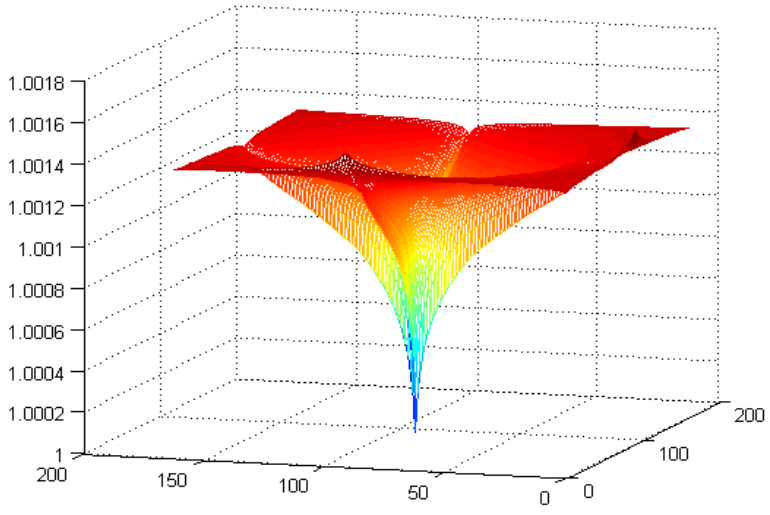


Figure 6.12 Inverted correlation plane after the application of non-linearity

In order to align the correlation plane and to minimize the energy levels at the edges the non-linearity function is applied using [0.1 0] as the parameters for equation 6.1. The resulting output is shown by Figure 6.13a below.

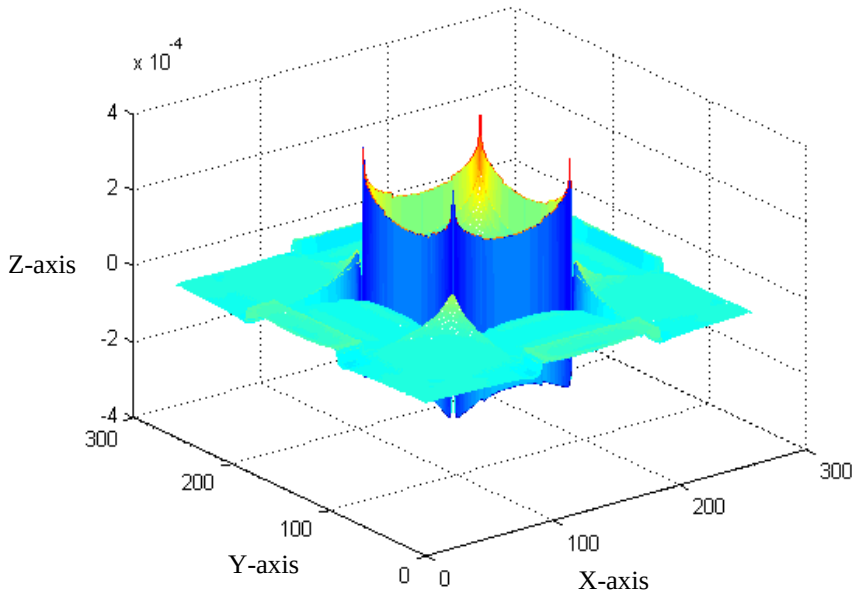


Figure 6.13a SPOT-MACH impulse response after application a non-linearity with a, c parameters [0.1 0]

The correlation output for the autocorrelation of the illuminated square after the application of non-linearity is shown by Figure 6.13b.

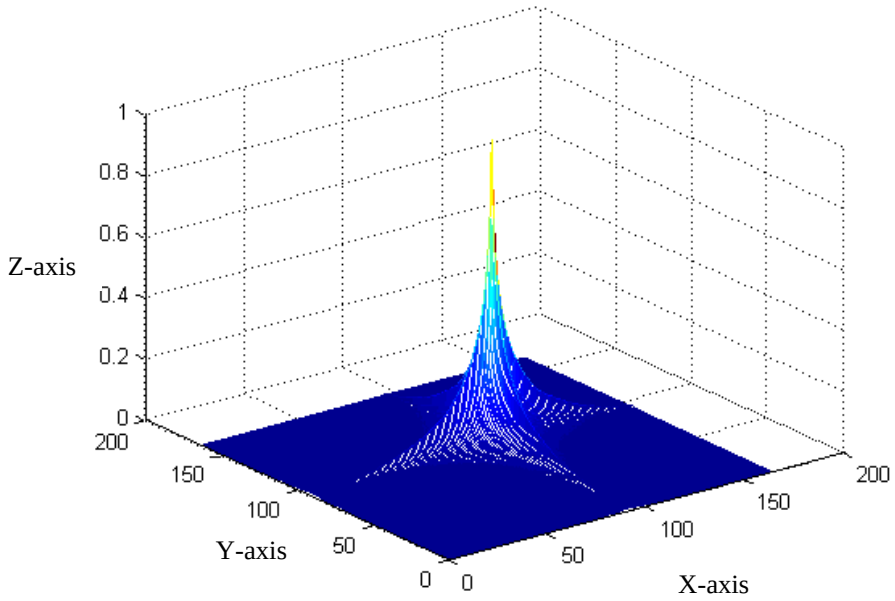


Figure 6.13b Correlation plane after application of non-linearity with a, c parameters [0.1 0]

It can be shown by comparing Figure 6.10(b) and Figure 6.13b that there is a reduction in the energy locked at the edges. Initially the energy at the edge was 4×10^{-3} which was reduced to 1.7×10^{-4} after the application of the non-linearity.

Based on this the non-linearity model with parameters [0.1 0] for equation 6.1 was applied to the transfer function of SPOT-MACH being used for FLIR imagery to check any improvement in performance assessed.

When the non-linearity was applied to the normalized kernel for the detection of multiple targets, due to nature of FLIR images where edges are smoothed by default and features are minimized, only a minimal increase in performance was seen. This is summarised in Table 6.7 given below.

FLIR Image (NP)	PCE	PSR	COPI
Without Non-Linearity	0.0016	15.35	0.0274
With Non-Linearity	0.0017	46.5	0.04

TABLE 6.7 Comparison of performance ratios with application of non-linearity

The application of non-linearity in spatial domain offers little advantage when considered for the case of FLIR images. But its application in terms of non-uniform brightness for frequency domain implementations is quite beneficial due to the fact that when an area is brightly illuminated it would have higher energy and the application of non-linearity would to smooth down the energy distribution to some extent [85].

6.2.2 Hardware Applications for Security Detection

For the application of the correlation technique to FLIR sensor imagery for security applications the only drawback of the spatial domain implementation is the significant computational requirement (effectively a 256x256 pixel size convolution kernel applied to a full resolution input image) [83].

A possible means of overcoming the computational requirements would be to employ the volume holographic based correlator system recently proposed by Birch et al [79]. Another alternative implementation which is better suited for the currently proposed application would be a FPGA based hardware correlator as proposed in Chapter 5.

The FPGA based correlator uses configurable logic to update the parametric values of the OT-MACH filter depending on the *COPI* values of the correlation plane. With the constant improvements in the processing speeds of these devices and built-in look up tables an efficient spatial domain convolution processor could be built [88].

An outline is given in Figure 6.14 for a proposed security application for perimeter protection especially border security, as wide areas have to be covered and considering the range on the FLIR sensor makes them an ideal choice.

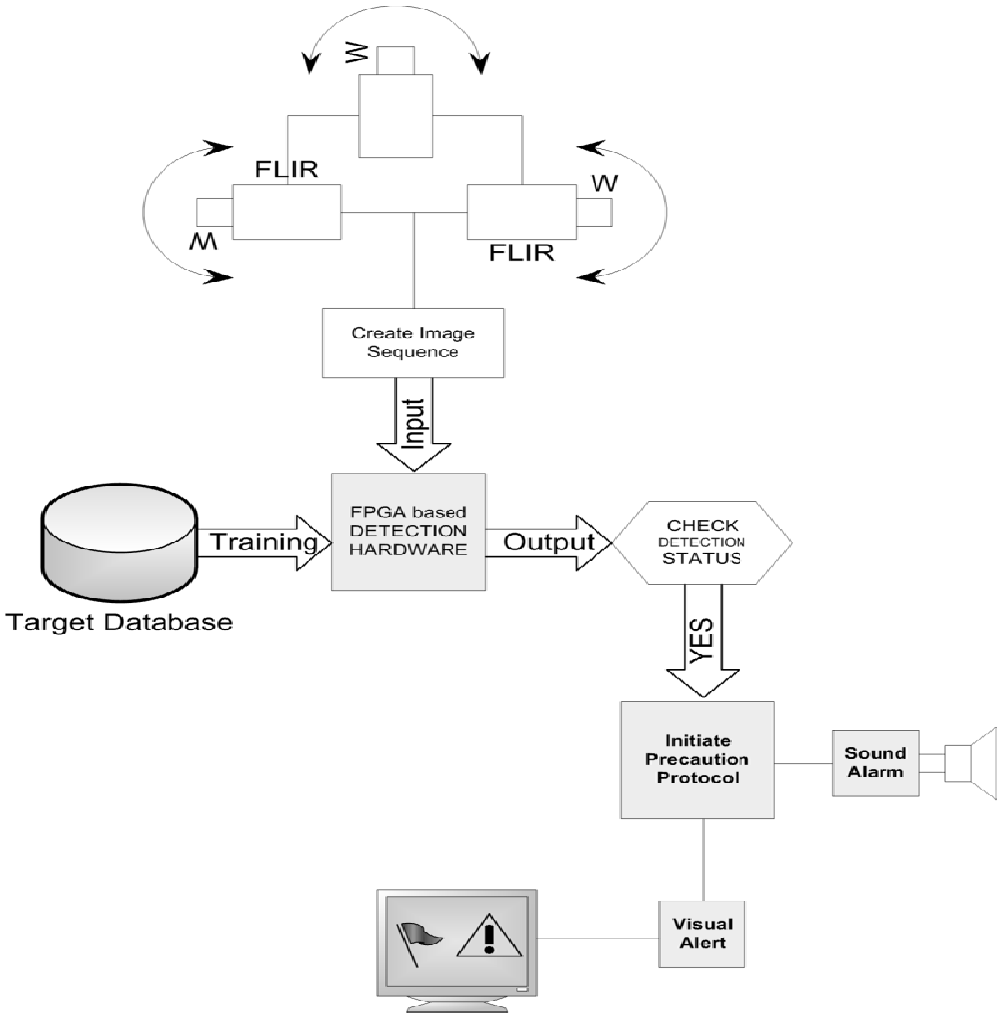


Figure 6.14 Proposed security application employing FPGA based correlator

Figure 6.14 is a perimeter monitoring system covering multiple degrees to avoid minimal blind spots for intruders connected to a sequencing system which converts the feed to sequence of time stamped images in raw format. The target database contains multiple classes of training images in this case a NP but there can be more than one trusted classes.

The FPGA based detection hardware using a SPOT-MACH filter is used to constantly correlate input image acquired from the FLIR sensors with the training image data set. In the case an object is detected which is not in the authorized class a security protocol is initiated.

This application can be tailored to operate in different scenarios or can be coupled as a part of a wider security system. Considering the proposed application there are still many areas which could be enhanced and modified with the rapid change in technology. In future work a more robust and faster system could be designed with additional features.

6.3 Overview of Feature Based Techniques

In order to extensively compare the proposed SPOT-MACH filter a very robust and efficient feature based technique known as Scale-Invariant Feature Transform (SIFT) proposed and patented by Lowe et.al has been selected [89]. But it should be noted that there is no direct comparison between correlation filters and Feature Based Techniques as correlation filters employ a top down approach whereas the later uses a bottom up approach. So in order to better understand the comparison results it is essential to understand the design and concept of SIFT [89] , [90].

The SIFT algorithm is based upon a concept that combines the scale-space theory and feature detection. In SIFT it is deduced that for any target in an image a group of interest points could be extracted providing a feature based description of the target. The extracted feature description from the image can be subsequently used to detect the target in a scenario where multiple objects are present in the scene. In order to perform a reliable detection the features extracted from the target scene should be invariant to changes in scale plus background noise and change in illumination, normally such group of features are present within the high contrast regions of the image which are located mostly on sharp edges [91].

The SIFT algorithm has been proven to be a very robust technique which can identify objects from cluttered backgrounds. This is due to the fact that the feature descriptors used by SIFT are tolerant to changes in scaling, orientation and partially invariant to distortion and illumination changes [89]. A more detailed working and implementation of SIFT has been discussed in the forthcoming section of this chapter.

The SIFT algorithm is based on the principle that the key points of the objects are first extracted from a set of reference images and stored in an image matrix. The key locations are defined as the maxima and minima of the DOG function applied in scale space to a series of smoothed and re-sampled images. The recognition process constitutes of matching each feature from the reference image to the input image using the features stored in the image matrix. A match is generated based on the Euclidean distance of each objects feature vectors. Using the full set of matches subsets of key points that agree on the target object and its location, scale and orientation in the input image are identified to filter out closest matches. The determination of matching clusters is effectively done by using a hash table implementation of the generalized Hough Transform. A cluster of three or more features that agree on a target object and its orientation are further subjected to detailed model verification and ultimately the outliers are discarded. In the end the probability that a set of features correspond to the presence of the target object has to be computed based on the accuracy of fit and number of false matches. The target object that passes all of these tests can be declared a match having a relatively high threshold [91].

The procedure discussed for the SIFT algorithm above has been divided into four key stages for the computation of feature vectors. These four key stages have been briefly described below. The detailed design and mathematical model of the SIFT can be understood by reviewing Lowe's paper [89].

6.3.1 Scale-Space Extrema Detection

For the implementation of SIFT the first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation [89].

6.3.2 Key Point Localization

After the selection of interest points at each candidate location, a detailed model is mapped to determine location and scale.The keypoints are selected based on measures of their stability [89].

6.3.3 Orientation Assignment

In SIFT one or more orientations are assigned to each key point location after the key points have been identified. This assignment is based on the direction of the gradient in the local image. All future operations are performed on the image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations [89].

6.3.4 Key Point Descriptor

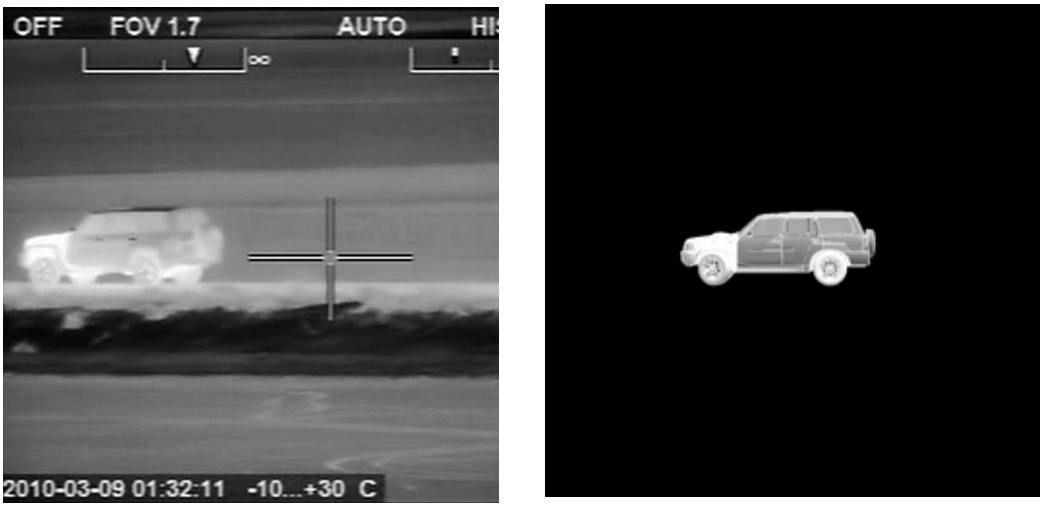
The local image gradients are measured at the selected scale in the region around each key point. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination. This makes SIFT not only invariant to change in orientations and scale but also to change in illumination [89].

6.4 Comparison of SPOT-MACH and SIFT using FLIR imagery

As discussed above that the SIFT technique is patented by Lowe so for the scope of this thesis the use of SIFT was not possible but another technique was used which was designed to produce results compatible to Lowe's version. The technique used is proposed by Vedaldi et.al. and gives 99.9% similar results to SIFT [50] , [36] and [92].

In this section the capabilities of SIFT has been compared with SPOT-MACH as FLIR images are intensity images and there is a lot variation in terms of heat signatures as the data set used is of harsh desert terrain an interesting set of results were obtained.

The following FLIR image was used as the target image and a 3D model of NP oriented at 180 degrees was used as the reference image as shown in Figure 6.15.



(a) **(b)**

Figure 6.15 Target Image for FLIR dataset (a), Reference Image for FLIR (b)

It was expected that SIFT being invariant to scale, orientations and illumination would give better results than the SPOT-MACH being invariant to illumination as well. The frames and descriptors were plotted for the SIFT first which is shown in Figure 6.16.

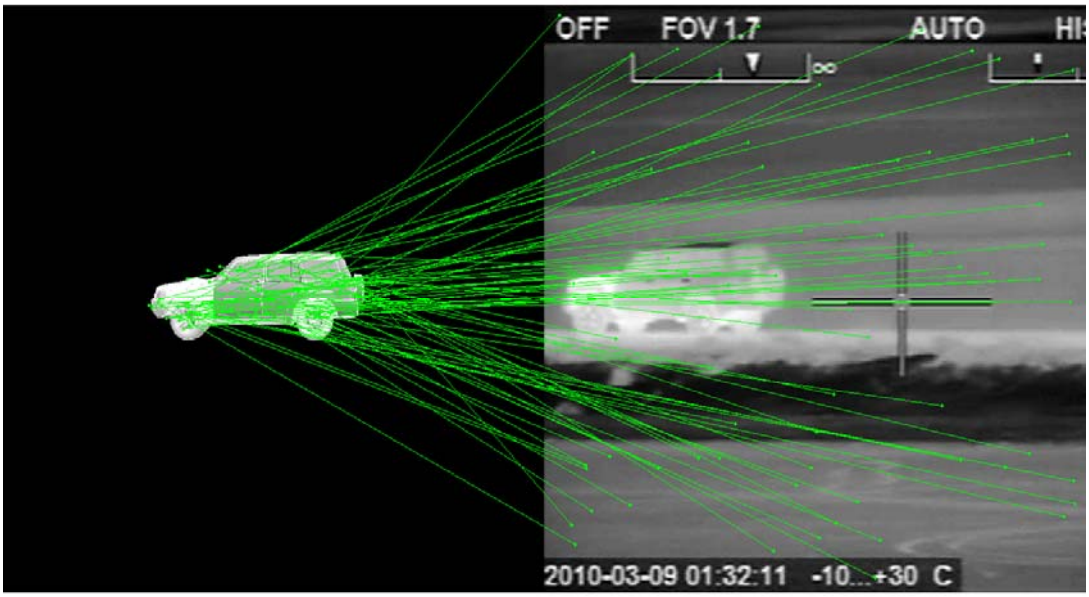


Figure 6.16 SIFT detection results for FLIR imagery

It can be seen from Figure 6.16 that none of the descriptors are matching the target image. There are a lot of false detections from SIFT in this specific scenario when using FLIR images. This could be due to the fact that the detection threshold of the SIFT might be set to behave as a low pass filter hence the number of matches, in order to improve the detection capabilities the threshold of detection is set higher to enable more refined matching which is shown in Figure 6.17.

Figure 6.17 SIFT detection results with increased threshold for FLIR imagery

From Figure 6.17 it is shown that when the detection threshold of the SIFT is increased even in this case only two descriptors are a match with the NP. And it has been discussed above that in order for a detection the SIFT should return at least 3-4 matches.

This is due to the fact that although SIFT is a very robust and invariant technique to orientations, scale and illumination it relies heavily on the quality of the image and fails where the image quality is low [93]. In the case of FLIR images there is not much high spatial frequency detail present in the image, so SIFT cannot generate an accurate match. In order to perform a comparison with the SPOT-MACH filter the same image was correlated using this filter and the following results achieved.

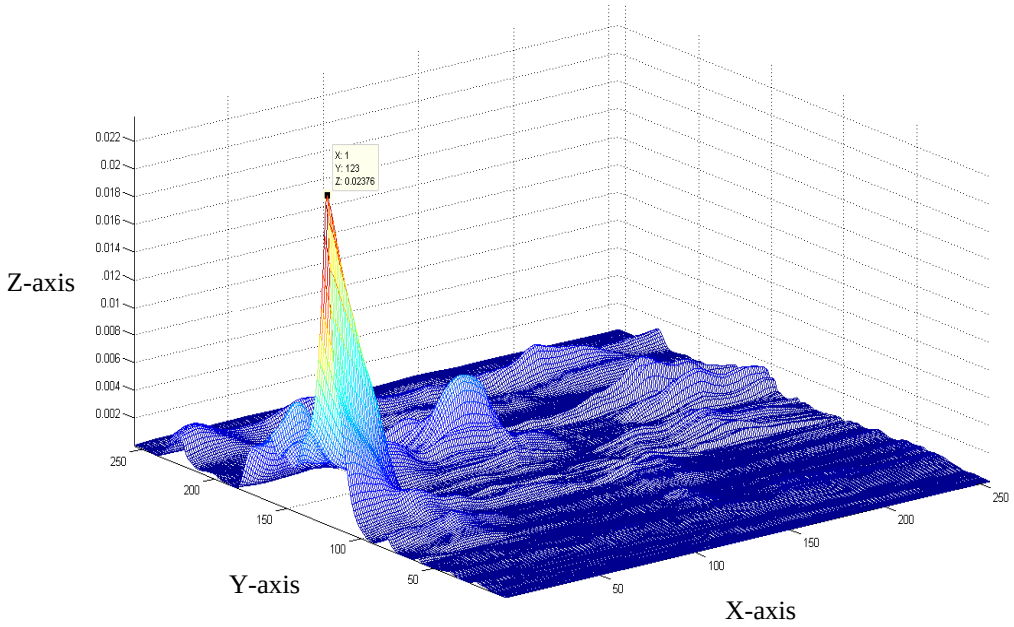


Figure 6.18 Correlation Output plane for SPOT-MACH using FLIR imagery

From Figure 6.18 the correlation output for SPOT-MACH is shown from which it can be seen that the SPOT-MACH successfully detects the target in the presence of noise and varying illuminations.

6.5 Comparison of SPOT-MACH and SIFT using visible imagery

In the previous section 6.3.2 the capabilities of SIFT were compared with SPOT-MACH and it was seen that the SIFT fails in situations where the image is not of high resolution and many variations are present in the background. In order to test the capabilities of SIFT and SPOT-MACH a testing library was created using a car as the target object and varying illumination profiles as background were added. The dataset used consists of a target image in which the target was a car oriented at 10 degrees.

In total five illumination test cases were prepared which are labeled as follows:

1. Uniform Lighting Test Case
2. Bright Illumination Test Case
3. Shadow Effect Test Case
4. Spot-Light Test Case
5. Dark Effect Test Case

Each of these test cases have been tested using SIFT, Frequency domain OT-MACH and SPOT-MACH to develop a performance comparison between these approaches.

6.5.1 Uniform Lighting Test Case

In the case of uniform lighting there was a single Omni-directional light source having no shadows with a uniform distribution of energy across the plane. In this case there were two scenarios that were created one with target object and another one with target and an up scaled false target. The test case created is shown in Figure 6.19.

(a) **(b)**

Figure 6.19 Uniform Lighting Test Case single target (a), Uniform Lighting Test Case False target (b)

The test results for Uniform Light test case given by Figure 6.19 are presented below.

I. SIFT results for Uniform Lighting Test Case

The test case was tested with SIFT algorithm and the following results were obtained for Figure 6.19(a).

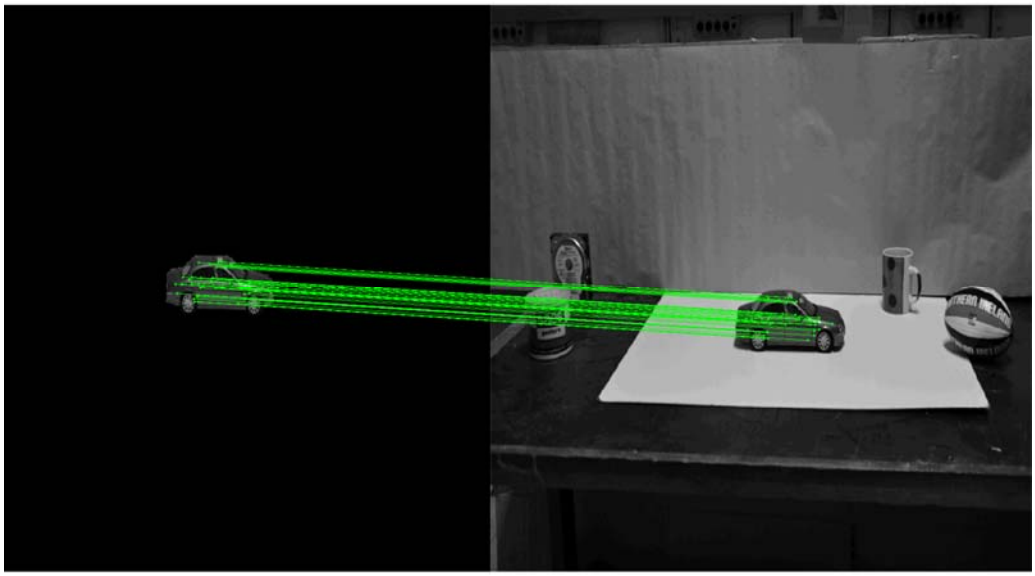


Figure 6.20 SIFT results for Uniform Lighting Test Case for image 6.19(a)

The test case was tested with SIFT algorithm and the following results were obtained for Figure 6.19(b).

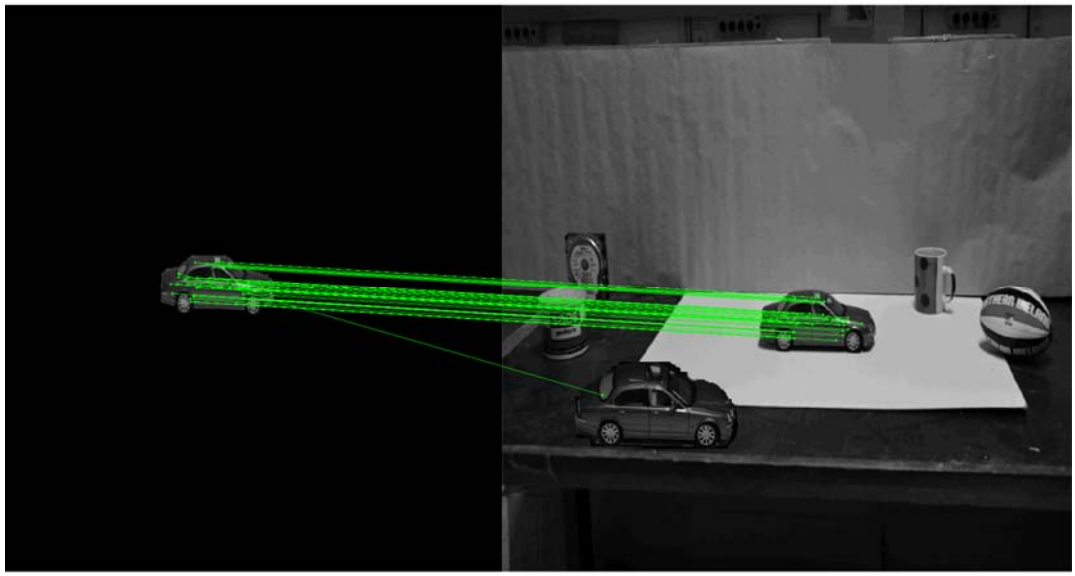


Figure 6.21 SIFT results for Uniform Lighting Test Case for image 6.19(b)

II. Frequency Domain OT-MACH results for Uniform Lighting Test Case

The Frequency domain OT-MACH has been tested using the car image oriented at 10 degrees as the reference image and Figure 6.19(a) as the target image. The same orientation reference image was used to maximse the chances of detection.

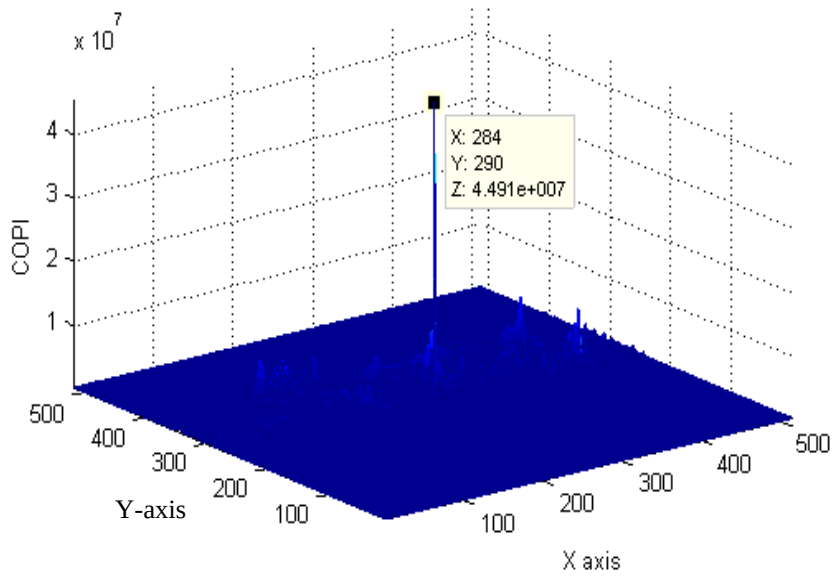


Figure 6.22 Output plane for frequency domain OT-MACH for single target

The frequency domain OT-MACH has also been tested for the image given in Figure 6.19(b).

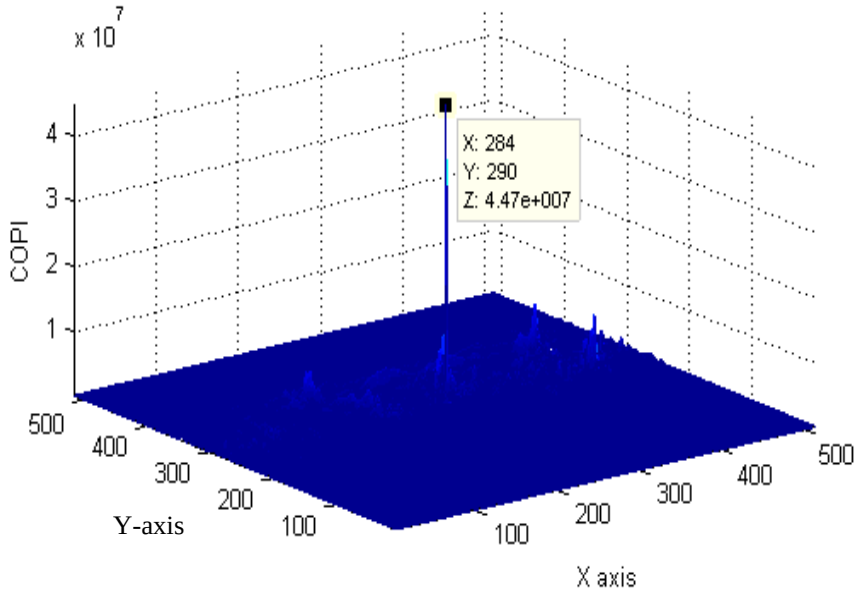


Figure 6.23 Output plane for frequency domain OT-MACH for two targets

I. SPOT-MACH results for Uniform Lighting Test Case

The SPOT-MACH has been tested on the using the car image oriented at 10 degrees as the reference image and Figure 6.19(a) as the target image.

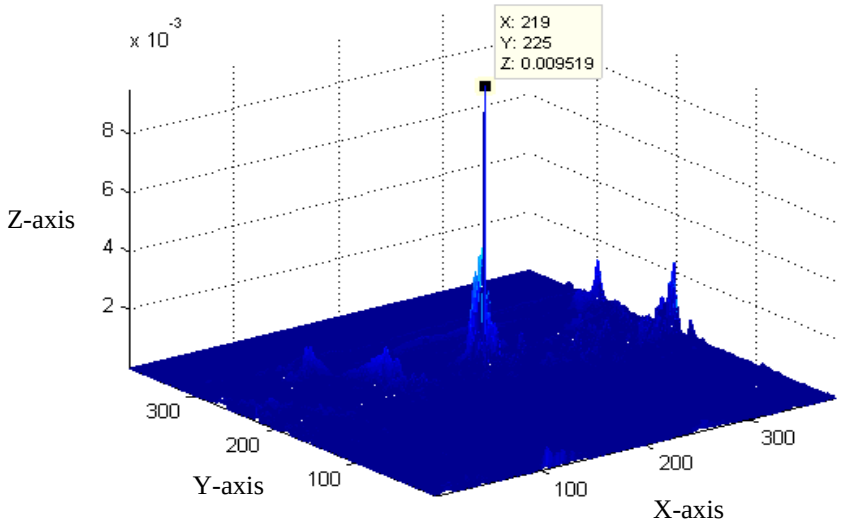


Figure 6.24 Correlation Output plane for SPOT-MACH for uniform lighting

The SPOT-MACH has also been tested for the Figure 6.19(b) containing a false target as shown by Figure 6.25.

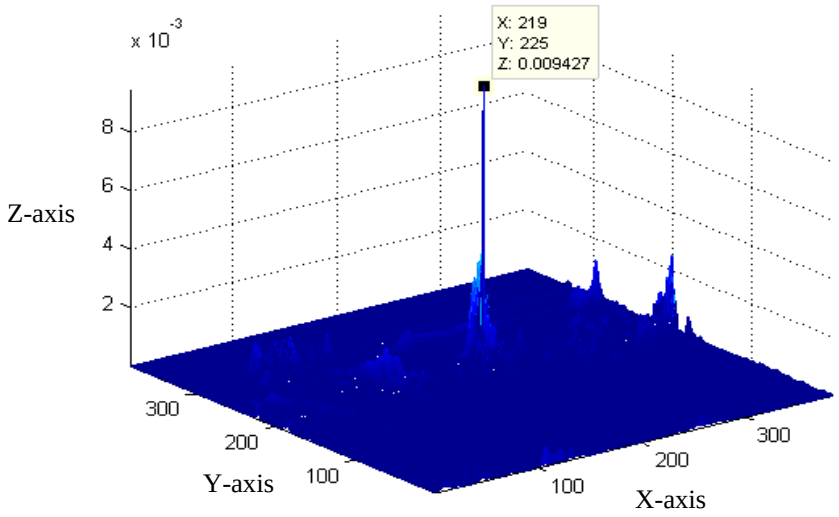
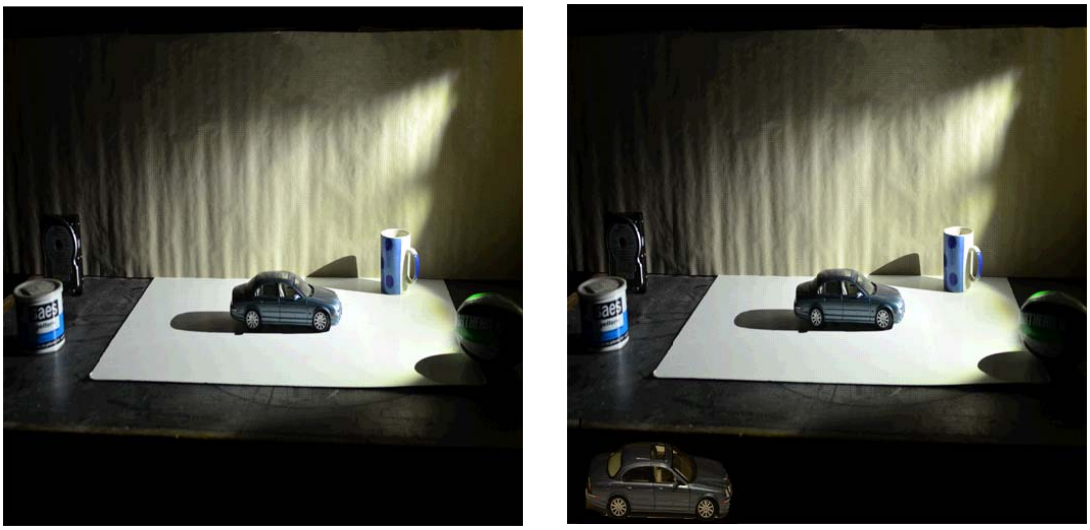


Figure 6.25 Correlation Output Plane for SPOT-MACH for uniform lighting with false target

6.5.2 Bright Illumination Test Case

In the case of Bright Illumination a directional light source has been used which illuminates the target object. In this case there were two scenarios that were created one with target object and another one with target and an up scaled false target in the non-illuminated area. The test case created can be shown from Figure 6.26.



(a) **(b)**

Figure 6.26 Bright Illumination Test Case single target (a), Bright Illumination Test Case False target (b)

The test results for Bright Illumination test case are presented below.

I. SIFT results for Bright Illumination Test Case

The test case was tested with SIFT algorithm and the following results were obtained for Figure 6.26(a).

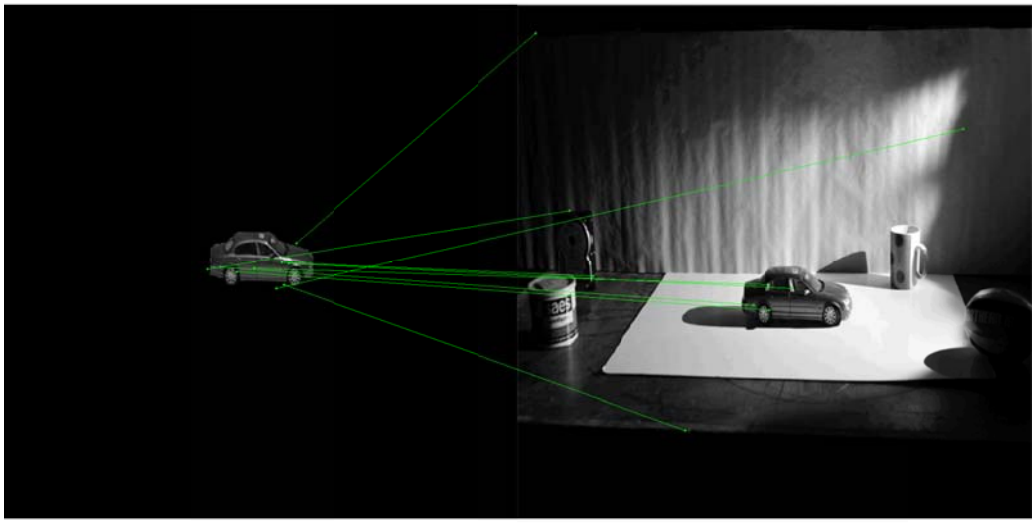


Figure 6.27 SIFT results for Bright Illumination Test Case for image 6.26(a)

The test case was tested with SIFT algorithm and the following results were obtained for Figure 6.26(b).

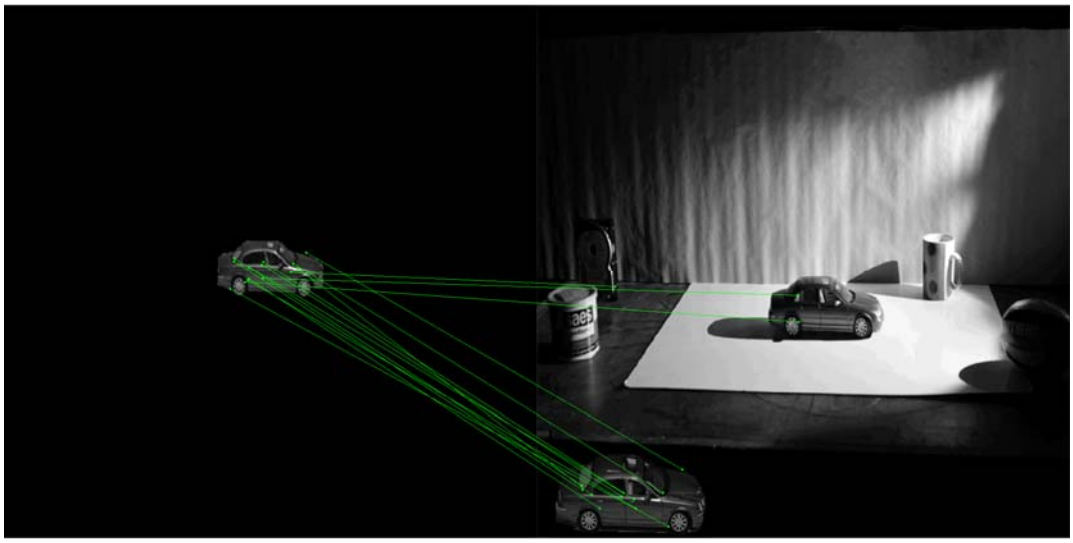


Figure 6.28 SIFT results for Bright Illumination Test Case for image 6.26(b)

II. Frequency Domain OT-MACH results for Bright Illumination Test Case

The Frequency domain OT-MACH has been tested using the car image oriented at 10 degrees as the reference image and Figure 6.26(a) as the target image.

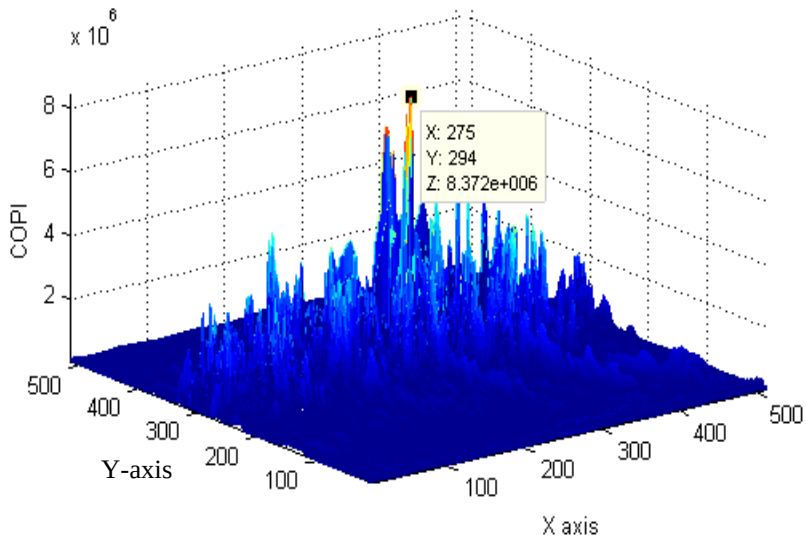


Figure 6.29 Output plane for frequency domain OT-MACH for single target in Bright Illumination Test Case

The frequency domain OT-MACH has also been tested for the image given in Figure 6.26(b).

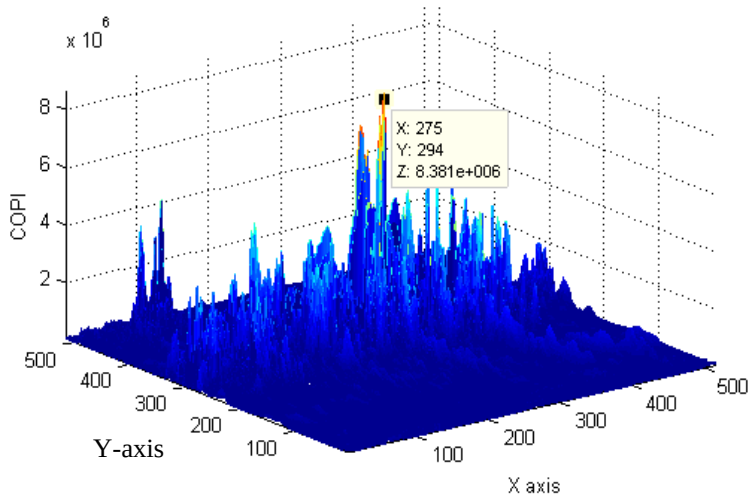


Figure 6.30 Output plane for frequency domain OT-MACH for two targets in Bright Illumination Test Case

It can be seen from Figure 6.29 and 6.30 that correlation was not successful in this case.

III. SPOT-MACH results for Bright Illumination Test Case

The SPOT-MACH has been tested on the using the car image oriented at 10 degrees as the reference image and Figure 6.26(a) as the target image.

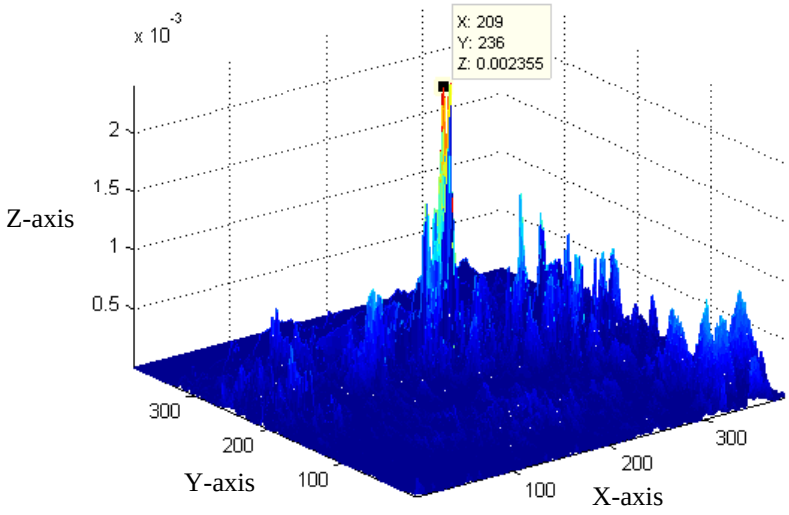


Figure 6.31 Correlation Output plane for SPOT-MACH for Bright Illumination

The SPOT-MACH has also been tested for the Figure 6.26(b) containing a false target as shown by Figure 6.32.

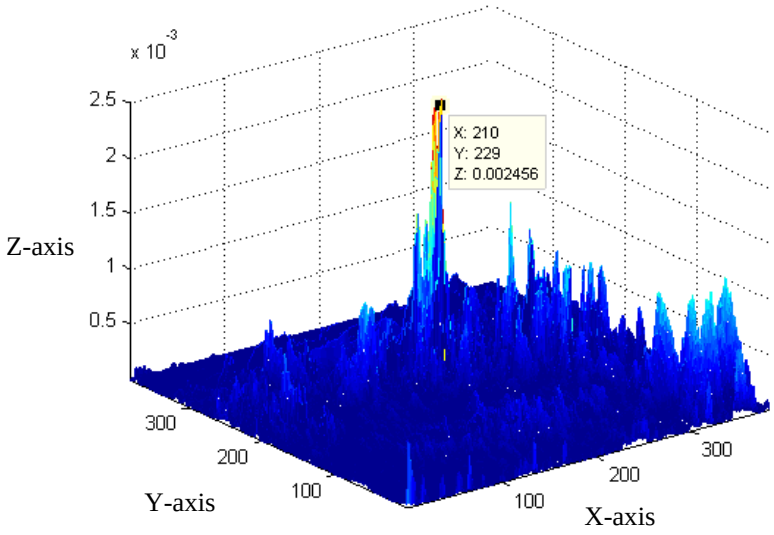


Figure 6.32 Correlation Output Plane for SPOT-MACH for Bright Illumination with false target

From Figure 6.32 it can be seen that the SPOT-MACH filter successfully detects the target object.

6.5.3 Shadow Effect Test Case

In the case of shadow effect test case an incandescent light source is used along with another light source targeted at the car from behind an obstacle to create a shadow effect. In this case there was a part of the car that was under the shadow of the obstacle, while one half of the full image was illuminated while the other half under a shadow. The test case created is shown from Figure 6.33.

Figure 6.33 Shadow Effect Test Case single target

The results for shadow effect test case given shown by Figure 6.33 are presented below.

I. SIFT results for Shadow Effect Test case

The test case was tested with SIFT algorithm and the following results were obtained for Figure 6.33.

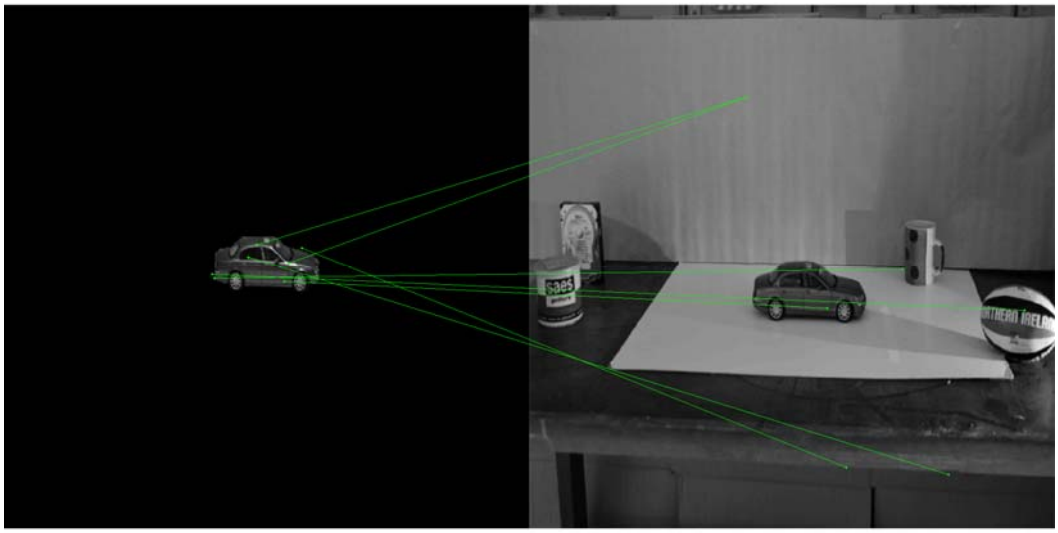


Figure 6.34 SIFT results for Shadow Effect Test Case

From Figure 6.34 it can be seen that the SIFT fails to give successful matches.

II. Frequency Domain OT-MACH results for Shadow Effect Test Case

The Frequency domain OT-MACH has been tested using the car image oriented at 10 degrees as the reference image and Figure 6.33 as the target image.

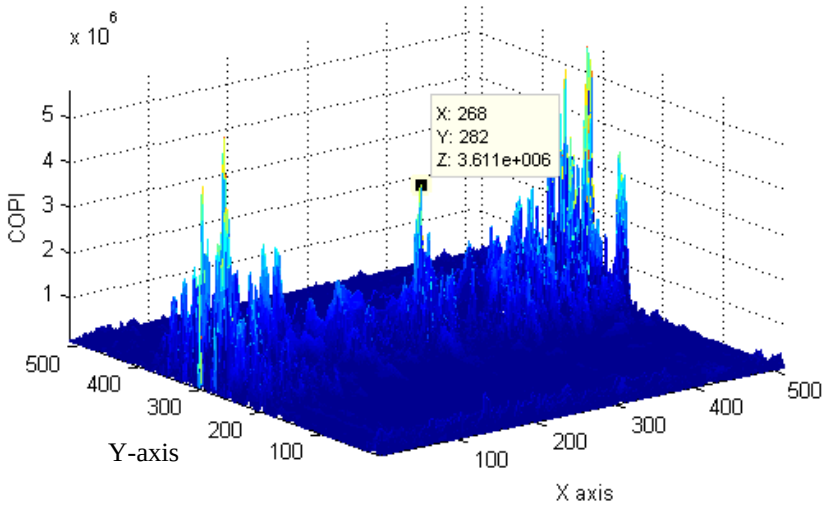


Figure 6.35 Output plane for frequency domain OT-MACH for Shadow Effect Test Case

From Figure 6.35 it can be seen that the OT-MACH gives a false detection.

III. SPOT-MACH results Shadow Effect Test Case

The SPOT-MACH has been tested on the using the car image oriented at 10 degrees as the reference image and Figure 6.33 as the target image.

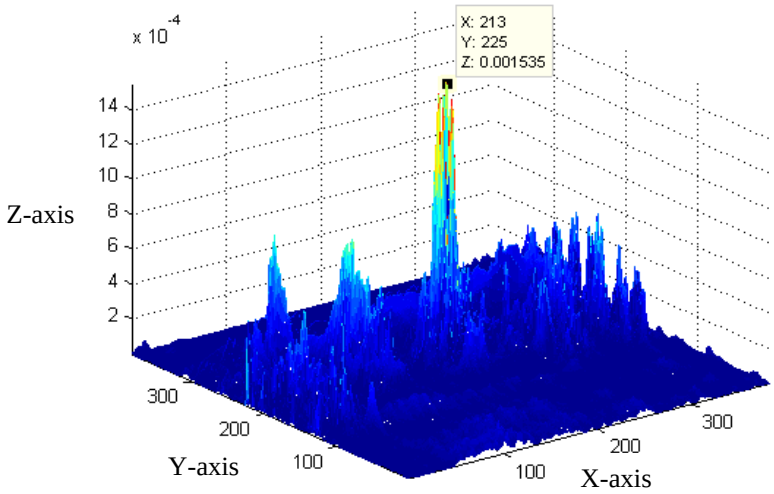


Figure 6.36 Correlation Output plane for SPOT-MACH for Shadow Effect Test Case

From Figure 6.36 it can be seen that the SPOT-MACH gives a peak at the location of the target object whereas the previous approaches discussed failed to do so.

6.5.4 Spot-Light Test Case

In the case of shadow effect test case a dimly lit incandescent light source is used along with another brightly illuminated light source targeted at the front part of the car creating a spot light effect. The Spot-Light test case created is shown from Figure 6.37.

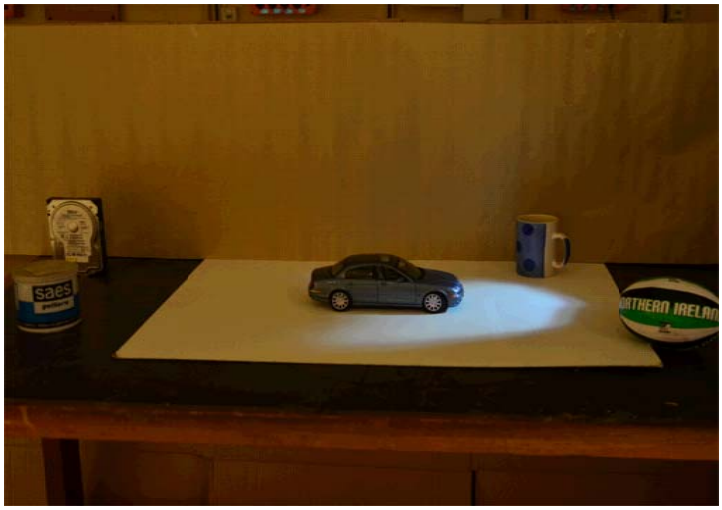


Figure 6.37 Spot-Light Test Case target image

The results for spot light effect test case given shown by Figure 6.37 are presented below.

I. SIFT results for Spot Light Test case

The test case was tested with SIFT algorithm and the following results were obtained for Figure 6.37.

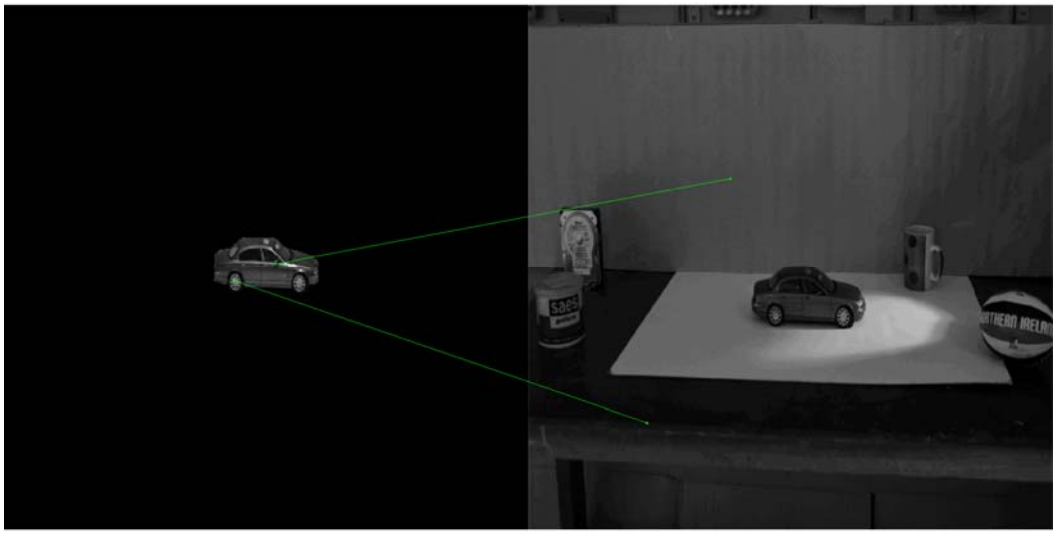


Figure 6.38 SIFT results for Spot Light Effect Test Case

From Figure 6.38 it can be seen that the SIFT clearly fails when there is a partly illuminated area present in the scene.

II. Frequency Domain OT-MACH results for Spot light Test Case

The Frequency domain OT-MACH has been tested using the car image oriented at 10 degrees as the reference image and Figure 6.37 as the target image.

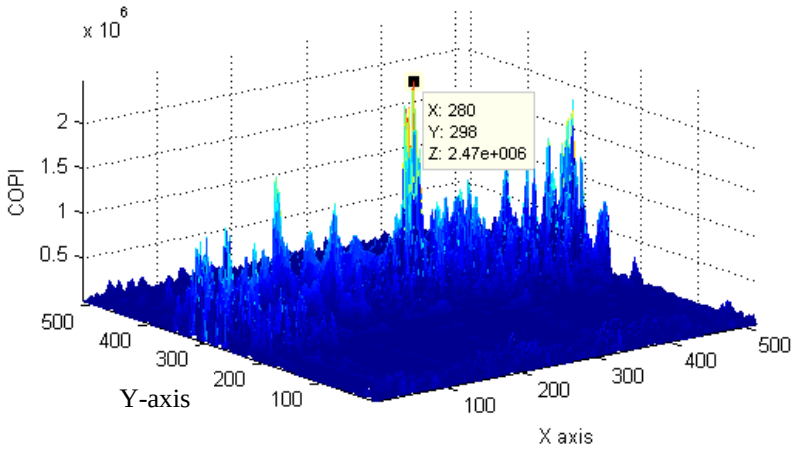


Figure 6.39 Output plane for frequency domain OT-MACH for Spot Light Test Case

From Figure 6.39 it can be seen that the OT-MACH produces a detection peak when there is a partly illuminated area present in the scene but also generates a false detection.

III. SPOT-MACH results for spot light effect test case

The SPOT-MACH has been tested on the using the car image oriented at 10 degrees as the reference image and Figure 6.37 as the target image.

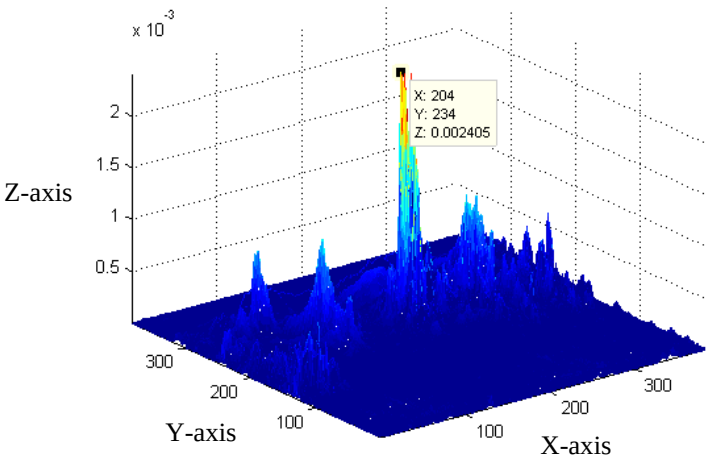


Figure 6.40 Correlation Output plane for SPOT-MACH for Spot Light Effect Test Case

From Figure 6.40 it can be seen that the SPOT-MACH produces a detection peak when there is a partly illuminated area present in the scene with effectively suppressing the background clutter to minimise false detections.

6.5.5 Dark Effect Test Case

The dark effect test case has been developed to test performance in terms of extreme illumination invariance. In this case whole of the car is in a dark area and bright light is shown on the wall away from the target to create an area of high energy to assess the performance of the techniques tested. The dark effect test case is shown from Figure 6.41.

Figure 6.41 Dark Effect Test Case target image

The results for dark effect test case given shown by Figure 6.41 are presented below.

I. SIFT results for Dark Effect Test case

The test case was tested with SIFT algorithm and the following results were obtained for Figure 6.41.

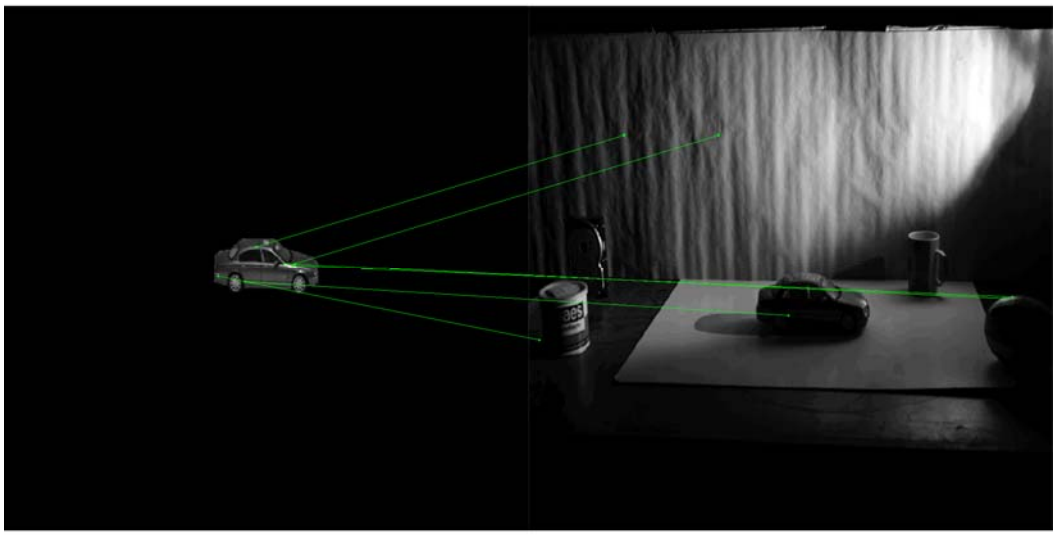


Figure 6.42 SIFT results for Dark Effect Test Case

From Figure 6.42 it can be seen that the SIFT produces only a single match in the absence of light which cannot be classified as a detection as three or more matches are required.

II. Frequency Domain OT-MACH results

The Frequency domain OT-MACH has been tested using the car image oriented at 10 degrees as the reference image and Figure 6.41 as the target image.

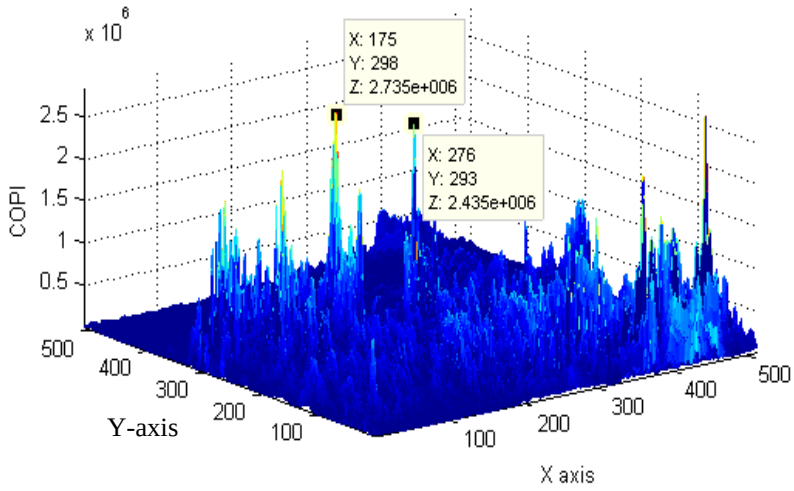


Figure 6.43 Output plane from frequency domain OT-MACH for Dark Effect Test Case

In the case of Dark Effect test case OT-MACH fails in the absence of a light source.

III. SPOT-MACH results for Dark Effect test case

The SPOT-MACH has been tested on the using the car image oriented at 10 degrees as the reference image and Figure 6.41 as the target image.

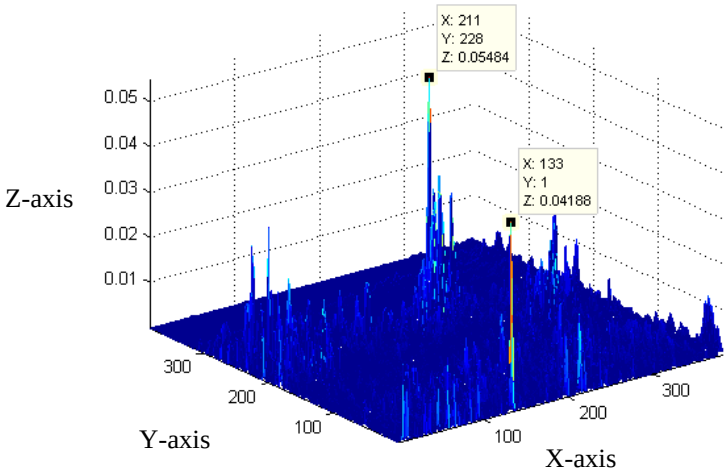


Figure 6.44 Correlation output plane for SPOT-MACH for Dark Effect Test Case

In the case of SPOT-MACH it can be seen that the location of the target object can be identified by the location of the peak.

6.6 Comparison Analysis

In Section 6.5 a detailed comparison of the SPOT-MACH filter was conducted with the Feature based technique and the frequency domain implementation of OT-MACH filter.

Based on the results from these tests, which were specifically developed to assess the performance of these techniques in terms of illumination invariance, a comparison matrix was created which is shown in Table 6.8 below.

Test Case	*OT-MACH*	*SIFT*	*SPOT-MACH*
Uniform Lighting	PASS	PASS	PASS
Bright Illumination	PASS	PASS	PASS
Shadow Effect	FAIL	FAIL	PASS
Spot Light	PASS	FAIL	PASS
Dark Effect	FAIL	FAIL	PASS

TABLE 6.8 Comparison Matrix for the OT-MACH, SIFT and SPOT-MACH filters

From Table 6.8 it can be seen that the SPOT-MACH filter offers the best overall performance in all the scenarios. The SIFT algorithm, although known to be capable of a degree of illumination invariance, was unable to generate matches in situations where the shadow was on the car or the spot light was illuminating a certain region of the car.

In the case of the Dark Effect Test Case, where the target car was completely in a region of very low illumination, only the SPOT-MACH filter was able to detect the exact position of the object in the scene.

This demonstrates that the application of local energy normalisation enables the SPOT-MACH filter to be invariant to changes in illumination. When combined with enhancement techniques, such as inclusion of a space domain non-linearity, the SPOT-MACH filter becomes an equally capable technique to the feature based techniques.

6.7 Conclusion

In this chapter FLIR images were used as a demanding test for the filters evaluated since they contain both high levels of variability of the thermal signature of the target object and high degrees clutter of the background desert scene within which it is contained. A security application was outlined which can be used for perimeter security and monitoring of targets in harsh desert conditions.

One of the main advantages of the SPOT-MACH filter is the ability to allow localised normalisation of the filter which is not possible in the frequency domain implementation of the filter. Results have been presented which indicate an orientation multiplexed spatial domain filter is able to detect, locate and recognise a target object within a cluttered scene. The filter was evaluated using the correlation plane metrics like PCE, PSR and COPI to quantify its detection capabilities within cluttered backgrounds. Also, a non-linearity was applied to the pre-filtered images and the effects of applying a sigmoid non-linear function were evaluated using the performance metrics.

Test cases were then created, using visual band imagery, containing varying degrees of illumination non-uniformity. These scenes were used for a detailed comparison of the SPOT-MACH filter performance with the SIFT technique and the frequency domain OT-MACH filter to test, in particular, the capability of the filters to detect and recognise a target object despite high degrees of illumination variation. The performance of the different methods was summarised in a Comparison Matrix from it is shown that the SPOT-MACH filter offers better detection and recognition performance than the other techniques under conditions of high illumination variation.

Chapter 7

Conclusions and future work

7.1 Conclusions

In this thesis a spatial domain correlation filter using local window energy normalisation was developed to overcome the problem of illumination variation in target images which adversely affects the performance of correlation filters. Illumination invariance is considered an integral criterion for the measurement of the performance of correlation filters. In this thesis a number of approaches were discussed to make a filter tolerant to illumination changes and it was concluded that frequency domain implementation techniques are inadequate for this requirement. However, the implementation of a spatial domain correlation filter came with the drawbacks of excessive computational requirements which were overcome by using speed enhancement techniques developed specifically for the spatial domain filter, which was named SPOT-MACH filter. Hence the SPOT-MACH filter is a two staged detection filter which utilizes enhancement techniques to make it comparable in terms of speed to its frequency domain counterparts.

In Chapter 1 it was shown that that correlation filters can be used to form an integral component of intelligent image pattern recognition (IPR) system which is capable of pre-processing data and defining a decision making criterion. For this it is necessary for a correlation filter to be invariant to all distortions within the target image whilst still being able to maintain a good discrimination between similar objects. So the filter required must establish a compromise between the conflicting requirements of in-class distortion tolerance and out-of-class discrimination; such a filter is called a distortion invariant filter. In the presence of the desired target the filter should also be able to determine its spatial location within the input scene.

Thus although computationally intensive, spatial domain correlation filters can use the method of normalized cross-correlation to match two objects.

The main advantage of using the normalized cross-correlation function is that this can be normalized for amplitude changes using the correlation coefficient in the spatial domain although this is a computationally intensive task. Due to the impracticalities involved in this process normalized cross-correlation was seldom used in the development of the initial correlation filters which, instead, employed the computational efficiency of the fast Fourier transform algorithm (FFT) to implement the filter in the frequency domain.

In Chapter 1 it was also shown that frequency domain correlation filters were employed to train the filter to be invariant to distortions of the target. The correlation coefficient cannot be implemented in the frequency domain hence making the frequency domain approaches un-normalized but with the trade-off of speed of implementation. One of the major drawbacks in the frequency domain distortion invariant filters discussed in this thesis was that false detections were given in the presence of non-uniform light distribution and hence the filters are not illumination invariant. This is due to the fact that global energy normalisation over the entire image will not work in scenarios where severe changes in lighting conditions are present. At the end of Chapter 1 a space variant approach to implement distortion invariant filters was proposed which offered to overcome the main drawbacks present in the frequency domain approaches.

In Chapter 2 the history and development of distortion invariant correlation filters was discussed. A brief overview of the SDF filter and its capabilities was given in terms of its frequency domain implementation. The MACE filter was initially the first SDF type filter to provide control over the shape of the output correlation plane by minimizing correlation energy using linear constraints. The limitations in SDF theory were discussed in terms of its inability to be invariant to the change in orientations of the target. The MACH filter was briefly discussed in this chapter and it was shown to have the ability to minimize distortions using a mean-squared-error criterion. It was concluded that the correlation filter design can offer better invariance to distortions in the presence of relaxed constraints as compared to hard constraints which has led to the design of unconstrained distortion invariant filters.

In Chapter 3 the MACH filter, which was briefly introduced in the previous chapter, was discussed in detail along with its enhancements. The MACH filter finds the optimal compromise between the pattern discrimination criteria in the presence of background noise and maximizes the relative height of the average correlation peak with respect to the expected distortions included in the design of the filter. A number of enhancements to the MACH filter were discussed and the OT-MACH filter was shown to be the preferred choice of implementation over its other enhancements like the EMACH filter. The frequency domain implementation of the OT-MACH filter was discussed in more detail and the discrimination abilities with and without background noise was tested using a visible imagery dataset. The results obtained from the visible imagery dataset for the frequency domain OT-MACH filter were compared in terms of performance ratios such as PCE and PSR. The test results which were obtained were used to show the degradation in performance of the frequency domain OT-MACH filter when additional orientations were added in to the filter. Also, the limitations of the OT-MACH filter and its inability to detect targets in non-uniform brightness were shown. An alternative spatial domain implementation was proposed which was named the SPOT-MACH filter.

In Chapter 4 a space variant approach to implement an unconstrained distortion invariant correlation filter was discussed. It was shown that the main advantage in the spatial domain was the ability to allow localized normalisation of the filter which is not possible in the frequency domain. The SPOT- MACH filter was created using a training image data set of visible imagery of a car in different orientations. In order to search for the desired target in the scene the SPOT-MACH filter was set up to search the entire target image step by step and so the correlation process was repeated for each location in the image. A robust moving window mechanism was used to overcome this which took a sub-image from the entire input image, its size being made equal to the window size which was set to be user dependant owing to the changing nature of target scenes.
It was decided that the size of the window should be small enough to reduce the simulation time and large enough to cover the distortions in the target object. A threshold was calculated by correlating each image with the SPOT-MACH filter and taking the average of their resultant correlation peak intensities. If the correlation peak values were above the threshold then the target was considered to be an in-class object where all the out-of-class objects values were lower than the threshold.

Also in Chapter 4 a comparison was performed with the frequency domain OT-MACH filter and it was shown that the performance of the OT-MACH filter is degraded with change of illumination. The SPOT-MACH filter was shown to overcome these problems by using local window energy normalisation but it was seen that the size of the window greatly influences the performance of the SPOT-MACH filter. The only disadvantage the spatial domain approach offers is the detection speed and in order to overcome this issue a set of enhancement of techniques were proposed for the SPOT-MACH filter, making it a two stage detection filter.

In Chapter 5 the enhancement techniques for the SPOT-MACH were discussed in order to make it comparable in speed of implementation to frequency domain approaches. Two methods of optimization were developed, the first employing software based techniques and the second using a hardware based approach. The first method employed low pass filtering the target image as a pre-processing stage to define potential regions of interests and then restricted the movement of the template window within these regions. A timing analysis was performed to show the reduction in processing time that was achieved because of this approach. A further software based approach that was developed was an entropy based filter which worked on the principle of background segmentation to define potential regions of interest where the target object might be located. This was considered to be a more controlled approach in terms of the accuracy of defining regions of interest where the object might be located as compared to the low-pass filtering approach.

There were two hardware based optimization approaches that were also proposed, using a FPGA and GPU to implement the SPOT-MACH filter (in the spatial domain).. The GPU based approach was tested using a number of window sizes ranging up to 1024x1024 within a 4096x4096 image. It was shown that the GPU based approach was effective only on larger window sizes where the performance gain was then noticeable. The FPGA based approach was then considered where the correlator was implemented on a SPARTAN II XC2S200e FPGA and a timing analysis was performed. A considerable gain in performance was observed due to the use of inbuilt multipliers and LUTs. Finally, it was concluded that the most optimal approach was to implement a hybrid optimization technique employing an entropy filter to define regions of interest and the FPGA based correlator to implement the matching operations..

In Chapter 6 a detailed performance analysis of the SPOT-MACH filter was performed using both visible and FLIR imagery and comparison made with the frequency domain implementation of and the SIFT algorithm. As mentioned in previously, one of the main advantages of the proposed SPOT-MACH is the ability to allow localized normalisation of the filter which is not possible in the frequency domain. In the comparative performance analysis, sample results were presented which indicated the SPOT-MACH filter's ability to recognize a target object within a cluttered scene. The performance of the filter was evaluated using the ratios PCE, PSR and COPI to determine the detection capabilities of the filter with cluttered backgrounds. Also, a non-linearity was applied to the pre-filtered images and the effects of the sigmoid non-linear function employed were evaluated using the performance ratios for the SPOT-MACH filter. It was observed that the application of a non-linearity in the spatial domain does not provide a significant performance gain with a FLIR dataset. Also in Chapter 6, FLIR images were used to assess the filters' performance in high levels of clutter and thermal signature variation within a scene. A security application was proposed which can be used for perimeter security and monitoring of targets in harsh desert conditions.

A number of test cases were then created in visible imagery based on varying illumination profiles. A detailed comparison was conducted of the SPOT-MACH filter with the SIFT algorithm and the frequency domain OT-MACH filter. A comparison matrix was created which showed that the SPOT-MACH filter offers better discrimination ability under conditions of severe illumination variability.

7.2 Future Work

In this thesis a spatial domain correlation filter, named the SPOT-MACH filter, was proposed which offers invariance to illumination as compared to its frequency domain equivalent. A lot of effort was put into the optimization of the SPOT-MACH filter for the implementation the correlation process in spatial domain. Although the optimization techniques proposed are effective, future implementations which are more robust and can be implemented quickly are needed with the continuing advancement in the size and pixel depth of images.

In manufacturing industry, Application Specific Integrated Circuits (ASIC) are used to form non-reconfigurable but extremely fast systems using silicon wafers from a specific die. However, these are expensive to produce and normally are profitable only when large scale production runs are required. Before anything can be taken to the manufacturing stage a prototype needs to be developed and tested. FPGA based Systems Bridge the gap by providing ASIC comparable speeds on a reconfigurable platform. FPGA applications have led to higher density devices; intellectual property (IP) integration and high-speed I/O interconnect technology. All of these elements have allowed FPGAs to play a central role in the defense and security industry.

A FPGA based computing board can be used to speed up computationally and data intensive algorithms. Before implementing the algorithm in hardware, the maximum speed-up gain for the desired algorithm must be considered. The quantitative estimate of the speed-up is dependent on the algorithm and the hardware computational resources available, although there is no general rule to estimate the speed-up gain for algorithms to be implemented on FPGA's. For future implementations, one proposed area is to develop the spatial correlator using a faster FPGA, as compared to the current SPARTAN XC2S200e.

For future implementation of the spatial correlator the defense-grade Virtex-5Q family is considered the best approach. This is because of the fact that it provides the newest, most capable features available in the aerospace and defense industries from the reprogrammable FPGA market leader.

Using the second generation Advanced Silicon Modular Block (ASMBL) column-based architecture, the Virtex-5Q family contains four distinct sub-families, the most offered by any FPGA vendor. Each sub-family contains a different ratio of features to address the needs of a wide variety of advanced designs. In addition to the most advanced, high-performance logic fabric, Virtex-5Q FPGAs contain many dedicated system-level blocks including powerful 36 Kbit block RAM/FIFOs and second generation 25 x 18 DSP slices [74] , [94].

The main reason for choosing the Virtex 5 family is due to the fact that it contains a 25 x 18 dedicated multiplier as well as an integrated adder for complex-multiply or

multiply-add operations. The dedicated large multiplier would offer a huge advantage in the computation of the correlation function in spatial domain [75], [95].

Another approach that can be considered for future implementation of the SPOT-MACH filter is the use of parallel computing or GRID computing to implement the computationally intensive correlation function [44] , [63] and [96].

Recently a lot of research has been done on the use of parallel computing for the implementation of detection and tracking algorithms. The SPOT-MACH correlator can be implemented using GRID computing techniques to employ the advantage of parallel processing along with mobility, as with the current advancement in wireless technologies, a detection system can be made employing wireless data transfer at speeds of up to 380Mbps to the GRID using internet connectivity, minimizing the need of hardware embedded within the sensor [97] , [88] and [98].

Thus with the ever increasing speed and flexibility of available hardware to implement the numerically intensive computations involved in spatial domain correlation based approaches, the accurate and high speed solution of difficult pattern recognition problems using this approach will become ever more feasible at reasonable cost [99] , [100].

References

1. **R. Duda, P. Hart.** *Pattern Classification and Scene Analysis* . s.l. : Wiley Interscience, 1973.
2. **T.Y. Young, K.S. Fu.** *Handbook of Pattern Recognition.* s.l. : Academic Press, 1986.
3. **A.Bhuiyan.** *Pattern Recognition and Tracking using Correlation Filters.* s.l. : VDM, 2008.
4. **Gilles Brassard, Paul Brately.** *Fundamentals of Algorithmics.* s.l. : Pearson Education, 2007.
5. **B.V.K. Vijay Kumar, A.Mahalanobis , R.D.Juday.** *Correlation Pattern Recognition.* s.l. : Cambridge University Press, 2005.
6. **L.F. Pau, C.H.Chen , P.S.P Wang.** *Handbook of Pattern Recognition and Computer Vision.* s.l. : World Scientific, 1993.
7. **R. Wu a, H. Stark.** *Rotation-invariant pattern recognition using a vector reference.* s.l. : Applied Optics, 1984. pp. 838-840. Vol. 23.
8. **C.F.Hester, D. Casasent.** *Multivariant technique for multiclass pattern recognition.* s.l. : Applied Optics, 1980. pp. 1758-1761. Vol. 19.
9. **A. Mahalonobis, B.V.K. Kumar , S. Song , S.R.F. Sims , J.F.Epperson.** *Unconstrained correlation filters.* 1994. pp. 3751-3759. Vol. 33.
10. **R.C. Gonzalez, R.E. Woods , S.L. Eddins.** *Digital Image Processing Using MATLAB.* s.l. : Pearson Education, 2006.
11. **R.C. Gonzalez, R.E. Woods.** *Digital Image Processing.* 2. s.l. : Prentice-Hall, 2002.
12. **A.Oppenheim, R.Schafer.** *Discrete-Time Signal Processing.* 2. s.l. : Pearson Education, 1999.
13. **Bow, Sing-Tze.** *Pattern Recognition: Applications to Large Data-Set Problems.* s.l. : MARCEK DEKKER, INC., 1984.
14. **S. Theodoridis, K. Koutroumbas.** *Pattern Recognition.* s.l. : Academic Press, 1990.
15. **M.Pavel.** *Fundamentals of Pattern Recognition.* s.l. : MARCEL DEKKER INC., 1989.
16. **A.Vanderlugt.** *Signal detection by complex spatial filtering.* s.l. : IEEE Transactions on Information Theory, 1964. pp. 139-146. Vol. 10.
17. **C. S. Weaver, J. W. Goodman.** *A technique for optically convolving two functions.* s.l. : Applied Optics, 1966. pp. 1248-1249. Vol. 5.
18. **B.Javidi, J.Wang.** *Optimum distortion invariant filters for detecting a noisy distorted target in background noise.* s.l. : Journal of Optical Society of America, 1984. pp. 2604-2614. Vol. 12.
19. **H.J.Caulfield, W.Maloney.** *Improved discrimination in optical character recognition.* s.l. : Applied Optics, 1969. pp. 2354-2356. Vol. 8.
20. **K.Morris, G. M.Mersereau.** *Scale, rotation and shift invariant image recognition.* s.l. : Applied Optics, 1986. pp. 2338-2342. Vol. 25.
21. **M. W. Farn, J. W. Goodman.** *Optimal maximum correlation filters for arbitrarily constrained devices.* s.l. : Applied Optics, 1989. pp. 3326-3366. Vol. 28.
22. **George.L.Turin.** *An introduction to Matched Filters.* s.l. : IRE Transactions on Information Theory, 1960. pp. 311-329. Vol. 6.

23. **J.G. Proakis, D.G. Manolakis.** *Digital Signal Processing.* 4. s.l. : Prentice-Hall, 2007.
24. **J. Thornton, M. Savvides , B.V.K Kumar.** *A Bayesian Approach to deformed Pattern Recognition Matching of Iris images.* s.l. : IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007. Vol. 29.
25. **F.T.S. Yu, X.J. Lu.** *A real time programmable joint transform correlator.* s.l. : Optical Communications, 1984. pp. 10-16. Vol. 52.
26. **U.Mahlab, J. Rosen , J. Shamir.** *Iterative generation of complex reference functions in a joint-transform correlator.* s.l. : Optics Letters, 1991. pp. 330-332. Vol. 16.
27. **B. Javidi, C.Kuo.** *Joint Transform Image correlation using a binary spatial light modulator at the Fourier plane.* s.l. : Applied Optics, 1988. pp. 663-665. Vol. 27.
28. **B.Javidi, J.Wang , Q.Tang.** *Multiple-object binary joint transform correlation using multiple-level threshold crossing.* s.l. : Applied Optics, 1991. pp. 4234-4244. Vol. 30.
29. **S. Jutamulia, D.A.Gregory.** *Soft blocking of the dc term in Fourier optical systems.* s.l. : Opt. Eng, 1998. pp. 49-51. Vol. 37.
30. **C. Li, S.Yin , F.T.S. Yu.** *Nonzero-order joint transform correlator.* s.l. : Optical Engineering, 1998. pp. 58-65. Vol. 37.
31. **G. Lu, Z. Zhang , S. Wu , F.T.S. Yu.** *Implementation of a non-zero order joint transform correlator by use of phase shifting techniques.* s.l. : Applied Optics, 1997. pp. 470-483. Vol. 36.
32. **D. Feng, H.Zhao , S.Xia.** *Amplitude-modulated JTC for improving correlation discrimination.* s.l. : Optical Communications, 1991. pp. 260-264. Vol. 86.
33. **H.J.Caulfield, R.Haimes.** *Generalized matched filtering.* s.l. : Applied Optics, 1980. pp. 181-183. Vol. 18.
34. **B.V.K.Vijaya Kumar, Wei Shi , C. Hendrix.** *Phase-only filters with maximally sharp correlation peaks.* s.l. : OPTICS LETTERS, 1990. pp. 807-809. Vol. 15.
35. **A. Strehl, J.K. Aggarwal.** *Detecting moving objects in airborne forward looking infra-red sequences.* s.l. : Journal of Machine Vision Applications, 2000. pp. 267-276. Vol. 11.
36. **A. Vedaldi, A. Zisserman.** *Self-similar sketch.* s.l. : Europeon Conference on Computer Vision(ECCV), 2012.
37. **D. Casasent, A. Furman.** *Sources of correlation degradation.* s.l. : Applied Optics, 1977. pp. 1652-1661. Vol. 23.
38. **A.Mahalanobis, R.Muise, S. R. Stanfill, , A. V. Nevel.** *Design and Application Of Quadratic Correlation Filters For Target Detection.* s.l. : IEEE Transactions Aerospace and Electonic Systems, 2004. pp. 837-850. Vol. 40.
39. **B.V.K. Kumar, L. Hasserbrook.** *Performance measures for correlation filters.* s.l. : Allied Optics, 1990. pp. 2997-3006. Vol. 29.
40. **A.A.S.Awwal, H.E.Michel.** *Enhancing the discrimination capability of Phase Only Filter.* s.l. : Asian Journal of Physics, 1999. pp. 381-384. Vol. 8.
41. **P.D.Gianino, J.L.Horner.** *Phase-only matched filtering.* s.l. : Applied Optics, 1984. pp. 812-816. Vol. 23.
42. **O.M.Parkhi, A. Vedaldi , A. Zisserman.** *On-the-fly specific person retrieval.* s.l. : Int. Workshop on Image Analysis for Multimedia Interactive Services, 2012.
43. **B. V. K. Vijaya Kumar, A. Lee, J. M. Connelly.** *Estimating object rotation and scale using correlation filters.* s.l. : Optical Engineering, 1989. pp. 474-481. Vol. 28.

44. **M.S.MILLAN, J.Campos , C.Ferreira , M.J.Yzuel.** *Matched filter and phase only filter performance in colour image recognition.* s.l. : Optics Communications. pp. 277-284. Vol. 73.
45. **Z.Bahri, B.V.K.Kumar.** *Generalized synthetic discriminant functions.* s.l. : Journal of Optical Society of America, 1988. pp. 562-571. Vol. 5.
46. **L.A. Romero, F.M. Dickey.** *Comparision between the peak-to-side lobe ratio of the matched and phase-only filters.* s.l. : OPTICS LETTERS, 1991. Vol. 16.
47. **D. Han, X. Li , R.N.Mohapatra , M. Michalak , Z. Nashed.** *Refining algorithms in correlation filter design for target detection.* s.l. : IEEE, 2008.
48. **A. Navel, A. Mahalanobis.** *Comparitive study of Maximum Aveage Correlation Height filtervariants using ladar remogery.* s.l. : Optical Engineering, 2004. Vol. 42.
49. **K. Brunndstorm, B.N. Schenkman , B. Jacobson.** *Object detection in cluttered infrared images.* s.l. : Optical Engineering, 2003. pp. 388-399. Vol. 42.
50. **A. Vedaldi, A. Zisserman.** *Sparse Kernel Approximations for Efficient Classification and Detection.* s.l. : IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2012.
51. **M. Savvides, B.V.K. Vijaya Kumar.** *Efficient design of advanced correlation filters for robust distortion-tolerant face recognition.* s.l. : IEEE, 2003. 0-7695-1971.
52. **R.A.Kerekes, B.V.K. Kumar.** *Selecting a composite correlation filter design:a survey and comapritve study.* s.l. : Optical Engineering, 2008. Vol. 47.
53. **K.Venkataramani, B.V.K. Kumar.** *Performance of composite correlation filters for fingerprint verification.* s.l. : Optical Engineering, 2004. pp. 1820-1827. Vol. 43.
54. **C.R.Chatwin, R. Wang.** *Frequency Domain Filtering Strategies for Hybrid Optical Information Processing.* s.l. : John Wiley and Sons, 1996.
55. **B.V.K Vijaya Kumar, D.W. Carlson , A.Mahalanobis.** *Optimal trade-off synthetic discriminant function filters for arbitrary devices.* s.l. : Optic Letters, 1994. pp. 1556-1558. Vol. 19.
56. **K.Al-Mashouq, B.V.K Vijaya Kumar.** *Analysis of signal-to-noise ratio of polynomial correltation filters.* s.l. : SPIE, 1999. pp. 407-413. Vol. 3715.
57. **E.Drexler.** *Nanosystems.* s.l. : John Wiley and Sons, 1992.
58. **P. Bone, R.C.D. Young , C.R. Chatwin.** *Position , rotation , scale and orientation invariant multple object recognition from cluttered scenes.* s.l. : Optical Engineering, 2006. Vol. 45.
59. **M.J.B.DUFF.** *Intermediate-Level Image Processing.* s.l. : Academic Press, 1986. ISBN 0-12-223325-5.
60. **S.Sims, J.Epperson , B.V.K.Kumar , A.Mahalanobis.** *SDF using relaxed constraints.* s.l. : SPIE. pp. 146-157. Vol. 1959.
61. **Yi.Li, Z.Wang , H.Zeng.** *Correlation Filter: An Accurate Approach to Detect and Locate Low Contrast Character Strings in Complex Table Environment.* s.l. : IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2004. pp. 1639-1644. Vol. 26.
62. **Kumar, B.V.K.V.** *Minimum variance synthetic discriminant function.* 1986. pp. 1579-1584. Vol. 3.
63. **L. Leclerc, Y. Sheng, , H. H. Arsenault.** *Rotation invariant phase-only and binary phase-only correlation.* s.l. : Applied Optics, 1989. pp. 1251-56. Vol. 28.
64. **B.V.K Kumar, A. Mahalanobis , S. Song , S.R.F.Sims , J. Epperson.** *Minimum Squared Error synthetic discriminant functions.* s.l. : Optical Engineering.
65. **A.Mahalonobis, B.V.K. Kumar , D. Casasent.** *Minimum average correlation energy filters.* s.l. : Applied Optics, 1987. pp. 3633-3630. Vol. 26.

66. **W. Dai, K. Wang.** *An Image Edge Dectection Algorithm Based on Local Entropy.* s.l. : IEEE, 2007.
67. **J.D.Brasher, M.Woodson.** *Composite training images for SDF's.* s.l. : Applied Optics, 1996. pp. 314-317. Vol. 35.
68. **A. Mahalonobis, B.V.K Kumar.** *Optimality of the maximum average correlation height filter for detection of target in noise.* s.l. : Optical Engineering, 1997. pp. 2642-2648. Vol. 36. 1.
69. **R.Hilai, J. Rubinstein.** *Recognition of rotated images by invariant Karhunen-Loeve expansion.* s.l. : J. Opt. Soc. Am. A, 1994. pp. 1610-1618. Vol. 11.
70. **P.Refregier.** *Optimal Trade-off Filters for Noise Robustness, Sharpness of the correlation peak and horner efficiency.* s.l. : Optics Letters. Vol. 16.
71. **Y. Sheng, H. H. Arsenault.** *Object detection from a real scene using the correlation peak coordinates of multiple circular harmonic filters.* s.l. : Applied Optics, 1989. pp. 245-249. Vol. 28.
72. **G. Ravichandaran, D. Casasent.** *Generalized in-plane rotation-invariant minimum average correlation energy filter.* s.l. : Optical Engineering, 1991. pp. 1601-1607. Vol. 30.
73. **R. Kerekes, B.V.K Kumar.** *Correlation filters with controlled scale response.* s.l. : IEEE Trans. Image Process., 2006. pp. 1794-1802. Vol. 15.
74. **M. Alkanhal, B.V.K.Kumar , A. Mahalanobis.** *Improving the false alarm capabilities of the maximum average correlaton height filter.* 2000. pp. 1133-1141. Vol. 39.
75. **Y.K.Seong, T.S.Choi.** *Optimal-trade-off filters for noise robustness , peak sharpness , and light efficiency in nonoverlapping background noise.* s.l. : Optical Engineering, 2000. pp. 45-52. Vol. 39.
76. **F. Goudail, P. Refregier.** *Optimal Detection of target with random gray levels on spatially disjoint background noise.* s.l. : Opt. Letters, 1996. pp. 495-497. Vol. 21.
77. **D.Casasent, W. Rozzi.** *Computer Generated and Phase-Only SDF Filters.* s.l. : Applied Optics, 1986. pp. 3767-3772. Vol. 25.
78. **A. Alkandri, A.Gardezi.** *Automatic parameter adjustment of difference of gaussian (DoG) filter to improve OT-MACH filter performance for target recognition applications.* s.l. : Proceedings of SPIE Defense and Security, 2011.
79. **P.M.Birch, A.A.Gardezi , B.K.Mitra , R. Young , C. Chatwin.** *An optical space domain volume holographic correlator.* s.l. : Proceedings of SPIE, 2009. Vol. 7340.
80. **O.Lepski, E.Mammen , G.V.Spokoinv.** *Optimal spatial adaptation to inhomogeneous smoothness: An approach based kernel estimates with variable bandwidth selectors.* s.l. : Annual Statist, 1997. pp. 929-947. Vol. 25.
81. **J.H.Park, G.S.Lee, W.H.Cho.** *Moving Object Detection Based on Clasius Entropy.* s.l. : IEEE, 2010.
82. **R.E.Walpole.** *Introduction to Statistics.* 3. s.l. : Macmillan, 1982.
83. **Ian Page, Michael Brady.** *Parallel Architectures and Computer Vision.* s.l. : Oxford Universiy Press, 1988.
84. **A. Gardezi, A. Alkandri ,P. Birch.** *A space variant MACH filter for object recognition in real time thermal images.* s.l. : Proceedings of SPIE, 2010. Vol. 7838.
85. **L.S. Jamal-Aldin, R.C.D Young , C.R.Chatwin.** *Application of nonlinearity to wavelet trasnformed images to improve correlation filter performance.* s.l. : Applied Optics, 1997. pp. 9212-9224. Vol. 36.
86. **A.Alkandri, A.Gardezi, P.Birch, R.Young, C.Chatwin.** *Parameter optimization of the optimal trade-off maximum average correlation height filter (OT-MACH) for FLIR imaging in high clutter environments.* s.l. : Proceedings of SPIE, 2011.

87. **M.N.Gibbs.** *IEEE Transactions on Neural Networks.* s.l. : IEEE, 2000. pp. 1458-1464. Vol. 11.
88. **A.E.Robinson, P. S. Hammon , V.R. de Sa.** *Explaining brightness illusions using spatial filtering and local response normalization.* s.l. : Elsevier, 2007. pp. 1631-1644. Vol. 47.
89. **D.Lowe.** *Distinctive Image Features from Scale-Invariant Keypoints.* 2004. pp. 91-110. Vol. 60.
90. **Y.Heo, K.Lee , S.Lee.** *Illumination and camera invariant stereo matching.* s.l. : Computer Vision and Pattern Recognition, 2008. pp. 1-8.
91. **L. Juan, O. Gwun.** *A comparison of SIFT , PCA-SIFT and SURF.* 2008. pp. 143-152. Vol. 3.
92. **K. Simonyan, A. Vedaldi , A. Zisserman.** *Descriptor Learning using Convex Optimisation.* s.l. : Europeon Conference on Computer Vision (ECCV), 2012.
93. **D. Devereux, B.K.Mitra , O. Holland.** *Using the Microsoft Kinect to model the environment of an anthropomimetic robot.* Pittsburgh : 2nd IASTED Intl. Conf. on Robotics, 2011.
94. **XILINX.** VIRTEX 5 Family. *XILINX website.* [Online] 2.0, March 22, 2010. [Cited: january 11, 2011.]
http://www.xilinx.com/support/documentation/data_sheets/ds174.pdf.
95. **S.Ali, S.Lucey.** *Are correlation filters useful for human action recognition?* s.l. : IEEE.
96. **M.Chen, Y. Li.** *An improved parallel FFT algorithm and application in Grid Environment.* s.l. : IEEE, 2010.
97. **A. Bakhtiari, A. Shirazi , A. Zahmati.** *An efficient segmentation method based on local entropy characterstics of Iris Biometrics.* s.l. : World Academy of Sciences, 2007. Vol. 28.
98. **Xi. Wang, Xin. Wang.** *FPGA Based parallel Architectures for Normalized Cross-Correlation.* s.l. : The 1st International Conference on Information Science and Engineering, 2009.
99. **B. Blakeslee, M.E.McCourt.** *A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization.* s.l. : Elsevier, 2004. pp. 2483-2503. Vol. 44.
100. **B.Blakeslee, M.McCourt.** *A multiscale spatial filtering account of Wertheimer-Benary effect and the corrugated Mondrian.* s.l. : Elsevier, 2001. pp. 2487-2502. Vol. 41.
101. **B.Blakeslee, M.E.McCourt.** *A multiscale spatial filtering account of the white effect, simultaneous brightness contrast and grating induction.* s.l. : Elsevier, 1999. pp. 4361-4377. Vol. 39.